SUBBAND CHANNELIZED RADAR DETECTION AND BANDWIDTH ESTIMATION FOR FPGA IMPLEMENTATION

BY

BRYAN THOMAS BURKE

B.S., United States Naval Academy, 1998

THESIS

Submitted in partial fulfillment of the requirements
for the degree of Master of Science in Electrical Engineering
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2000

Urbana, Illinois

SUBBAND CHANNELIZED RADAR DETECTION AND BANDWIDTH ESTIMATION FOR FPGA IMPLEMENTATION

BY

BRYAN THOMAS BURKE

B.S., United States Naval Academy, 1998

THESIS

Submitted in partial fulfillment of the requirements
for the degree of Master of Science in Electrical Engineering
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2000

Urbana, Illinois

ABSTRACT

The theory of optimum radar detection is well known and is generally implemented in expensive ASICs or supercomputers. However, today's state-of-the-art FPGAs are capable of performing relatively complex algorithms and provide the added benefit of being reconfigurable with new algorithms or methods on-site. Los Alamos National Laboratory has undertaken the goal of developing a receiver that is capable of performing detection and bandwidth estimation of pulsed radar systems. It is designed to function in electronic intelligence (ELINT) applications, where the goal is to determine the capabilities of threatening systems, such as radars which guide aircraft or missiles to targets.

This thesis addresses methods of pulse detection and bandwidth estimation that are able to be implemented on an FPGA. The framework is that which is commonly used in this application: a polyphase filter bank subband frequency decomposition of the RF signal, followed by statistical detection methods. The optimal fixed-sample-size (FSS) estimator for this subband decomposition is shown to be the F-test, based on the output statistics of the filter bank, which are found to be chi-squared. An alternative to fixed-sample-size methods, the sequential probability ratio test (SPRT) is, however, more suited to ELINT due to its ability to adapt the test length to the received data. The SPRT is shown to achieve a higher probability of detection with approximately 1/5 the required sample size of the FSS method. The complexity of the SPRT is equivalent to that of the FSS method, and the statistic that results from the optimal SPRT implementation also lends itself to easy calculation of the bandwidth of the signal.

ACKNOWLEDGMENTS

I would like to thank Dr. Joe Arrowood and Dr. Mark Dunham for the opportunity, help, and guidance in working on this project, as well as Professor Doug Jones for allowing me the freedom to finish this work.

I would also like to thank my family who have always supported me in my academics and have provided me with encouragement along the way. Finally, I would like to thank Miquela for the motivation to finish this thesis when I never thought it would be done.

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

ADC	Analog-to-digital converter
ANOVA	Analysis of variance
ASIC	Application-specific integrated circuit
ASN	Average sample number
CDF	Cumulative distribution function
CUSUM	Cumulative sum test
CW	Continuous wave
DCT	Discrete cosine transform
DFT	Discrete Fourier transform
ELINT	Electronic intelligence
FAR	False alarm rate
FPGA	Field programmable gate array
FSS	Fixed sample size
HDL	Hardware description language
i.i.d.	Independent and identically distributed
LFMOP	Linear frequency modulation on pulse
ML	Maximum likelihood
MOP	Modulation on pulse
NOMOP	No modulation on pulse
OCF	Operating characteristic function
PSKMOP	Phase shift keying modulation on pulse
pdf	Probability density function
RC	Reconfigurable computing
RF	Radio frequency

SNR	Signal-to-noise ratio
SPRT	Sequential probability ratio test
STFT	Short-time Fourier transform
TOA	Time of arrival
TOD	Time of departure
UMP	Uniformly most powerful

CHAPTER 1

INTRODUCTION

For years, the theory of optimum radar detection has been known and performed in practice using analog methods, expensive ASIC implementations, or supercomputers. The recent advances in FPGA technology have caused a new paradigm in algorithm development, the reconfigurable computing architecture. Algorithms coded in the hardware description language (HDL) can be loaded onto FPGAs to perform signal processing tasks at a speed that is not possible today in low-cost floating-point processors. The ability of the FPGA to be reprogrammed with a new or updated algorithm allows it a flexibility that is not available in custom-designed ASICs. With this reconfigurable flexibility come certain limitations, however. The computational complexity that is achievable in today's state-of-the-art FPGAs is a function of the number of logic components that can be fit on a chip. Also, the FPGA only functions efficiently in a fixed-point architecture, which does not lend itself to complex signal processing algorithms. As such, it requires a joint optimization of resources to reduce the computational complexity of the algorithms so they will perform well in the environment of reconfigurable computing, while maintaining a level of performance that justifies the use of RC technology over more expensive ASICs.

A leader in RC technology, Los Alamos National Laboratory's Space Engineering Division has recently developed an RC platform suitable for signal processing applications and has undertaken the goal of implementing a matched bandwidth radar receiver in FPGA technology. The design goal is to implement a computationally efficient radar pulse detection strategy, while maintaining as high a level of performance with respect to optimal methods as possible. This thesis work provides a piece of the overall design by investigating the best pulse detection strat-

egy followed by a bandwidth estimation scheme that can be implemented given the restrictions of the RC platform, the RCA-2.

1.1 Problem Statement

The particular design problem posed was that commonly incurred in an electronic intelligence (ELINT) application. ELINT is the result of observing signals transmitted by radar systems to obtain information about their capabilities [1]. The value of ELINT is that it provides information about threatening systems, such as radars that guide aircraft or missiles to targets. Clearly, ELINT is most useful in situations where some hostility is involved; otherwise the information could be obtained directly from the radar user or designer. The design problem is difficult in that, for ELINT applications, a priori knowledge of the systems to be detected is minimal. Besides the knowledge that the detector should find pulsed radar systems, a minimum of other assumptions should be made. The design method can be summarized most effectively as a series of steps:

- Precondition the signal, possibly involving linear or nonlinear transformations.
- Implement the "best" detector, given the computational complexity limitations of the host platform.
- Upon detection, deduce as much information regarding the detected event as possible (e.g., time of arrival (TOA), pulse length, SNR, bandwidth, center frequency, pulse modulation).

In fact, the optimal solution to this problem is known; however, the goal of this work is to see if the optimal solution can be implemented in an FPGA, and whether it is wise to do so. There may be assumptions and simplifications that allow near-optimal performance for much reduced computational complexity. As always, engineering design is a balance of trade-offs between performance and implementation issues.

1.2 Design Parameters

Several of the design parameters of the problem are limited by the RC platform itself and must be heeded during the design process, outlined in Table 1.1. Other parameters are assumptions based on the typical environment of this system. Most of the design parameters

Table 1.1 Design Parameters

Parameter	Value	Unit
Sampling Rate	100	MHz
Bandwidth	5 – 45	MHz
Pulse Duration	.1 – 1000	μs
Pulse SNR	-20 – 20	dB
Center Frequency Agility	1	μs
Modulation	NOMOP, LFMOP, PSKMOP, Hopped	
Noise	White, Gaussian, Non-stationary	

are self-explanatory; however, those having to do with the center frequency agility and the modulation will be dealt with fully in Chapter 5 and are peculiar to this application.

1.3 System Overview

As stated earlier, the design goal is a matched-bandwidth pulse detector. A matched-bandwidth pulse detector is one in which pulses are detected in the maximum system bandwidth but are then modulated to baseband and filtered with a filter matched to their modulation or pulse bandwidth in order to improve the resulting SNR of the pulse. After this process, there are many options, including compression, recording, or further processing of the enhanced waveform. At the system level, the process is a parallel one. We are allowed to manipulate the input RF data in any way, destructive or otherwise, to accomplish detection, as the same RF data is input into a digital modulator which translates the pulse center frequency to zero, then uses an adaptive decimation scheme to match the bandwidth of the pulse. Figure 1.1 presents a system block diagram. The individual components of the system will be discussed more fully, beginning with the polyphase filter bank in Chapter 2. Two options for statistical pulse detection are presented: the fixed sample size estimators in Chapter 3, and sequential estimators in Chapter 4. Following detection, bandwidth and center frequency must be deduced, and are the subjects of Chapter 5.

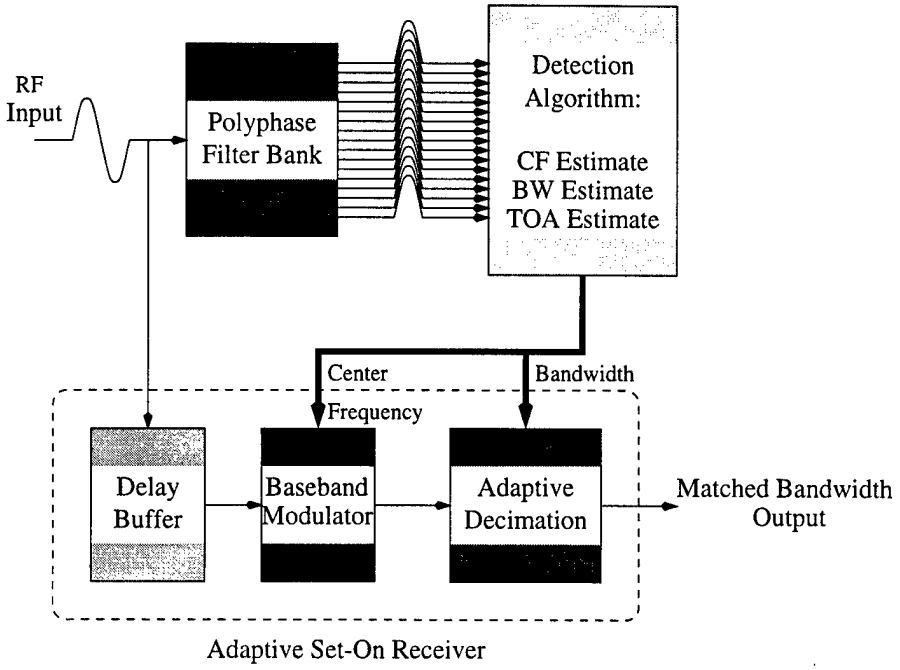


Figure 1.1 Matched bandwidth detector/receiver.

CHAPTER 2

POLYPHASE CHANNELIZER

The first step in the process of matched bandwidth detection is the subband channelization, which is performed to increase the SNR of pulses, as we expect them to have a finite frequency support limited to some ratio of the full input bandwidth. Of course, there are many ways that one can go about the filtering that produces a subband decomposition, but the theory rests on two operations: a prototype filter to achieve the desired subband shape, and the modulation method which produces the channels across some portion of the spectrum. In our case, there are three main options, a wavelet decomposition, a filter bank based on a discrete cosine transform (DCT) modulator, and a filter bank modulation based upon a discrete Fourier transform (DFT). For FPGA implementation, we choose to use a DFT-based scheme because there exist many cores for commonly used FPGAs to perform both complex and real-valued DFTs. Wavelet and DCT methods are possible, and are investigated fully in a paper by Arrowood [2].

Given that we are using a DFT as the modulator for our decomposition, a polyphase implementation makes the most sense and provides a multirate architecture that can allow the internal clock used for computation to be set at a much lower rate than the input data clock. A polyphase filter bank is a simple system, based on two steps; first, the signal is filtered by a prototype filter, and then the resulting signal is transformed via the DFT. A third step, which is important to the nature of this project, is what to do with the DFT output. The resulting complex output has twice as many data points as the input, because each input sample produces a real and imaginary part which must be accounted for in the FPGA individually. As such, one option is to use the magnitude of the output, thus resulting in a 1:1 data transformation.

Since the goal of our subband decomposition is detection, we must derive the expected statistics of the output of the polyphase filter bank in order to determine the optimal methods

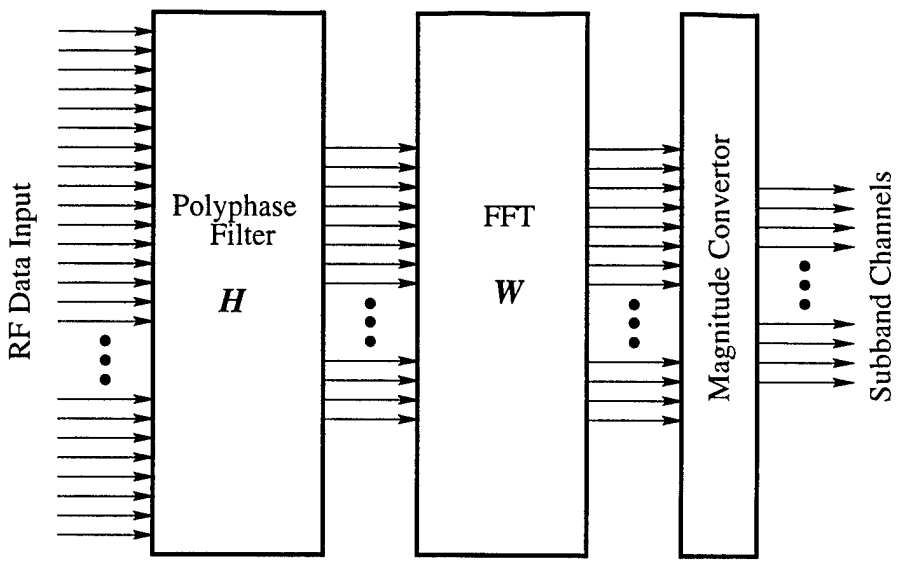


Figure 2.1 Polyphase filter bank.

of detection. To begin, there must be some assumptions made. We will assume the wideband RF noise to be Gaussian, white, and nonstationary. The nonstationarity is important for real-world applications, but in most derivations, we will assume that over a short time interval a random sample is independent and identically distributed (i.i.d.). Our method of derivation is simple. Model each process as a transformation, and apply methods of statistical derivation appropriate to each type of transform.

2.1 Reduced Signal Model

As a first approximation, let us propose a reduced signal model where the pulse is of finite time and frequency support, with no modulation either on the pulse envelope or frequency basis. Surface search radar employs this type of pulse often; however, it is far from the expected in the case of more sophisticated radars, as modulation provides important sensitivity increases. Let our received signal be of the form

$$r(k) = s(k) + n(k) \tag{2.1}$$

where $s(k)$ is the desired pulsed signal and $n(k)$ is the noise, subject to the assumptions above. Let us make several assumptions regarding the signal and noise that provide a reduced signal model but give us a roadmap for more complicated signal models. Specifically, we assume:

- The DFT length is N_f and the polyphase prototype filter is a boxcar filter of length $N_h = N_f$.
- The signal pulse is of length $N_w = N_f$.
- The signal is unmodulated either in pulse envelope or in frequency, but can contain some initial phase offset δ which is modeled as a uniform density on the interval $[0, 2\pi)$.
- The base frequency of the signal is centered in one channel of the decomposition.

Given the above assumptions, $s(k)$ is given by

$$s(k) = A\cos(\omega(k) + \delta) \tag{2.2}$$

where A is the amplitude, δ is the initial phase, and $w(k)$ is of the form

$$\omega(k) = e^{j\frac{2\pi kf}{N_f}} \tag{2.3}$$

In the absence of noise, the received signal is $r(k) = s(k)$. We model the DFT as a linear transformation given by

$$\mathbf{W} = \begin{bmatrix} W_{0,0} & W_{0,1} & \cdots & W_{0,N_f} \\ W_{1,0} & \ddots & & \\ \vdots & & \ddots & \\ W_{N_f,0} & & & W_{N_f,N_f} \end{bmatrix}$$

where $W_{l,k}$ is

$$W_{l,k} = e^{-j\frac{2\pi kl}{N_f}} \tag{2.4}$$

If we apply a sequence of linear transformations, $\mathbf{W}$ for the DFT and $\mathbf{H}$ for the polyphase filtering, the real and imaginary parts of the transformed signal are

$$\mathbf{X_s} = Re(\mathbf{WHs}) \tag{2.5}$$

and

$$\mathbf{Y_s} = Im(\mathbf{WHs}) \tag{2.6}$$

where $\mathbf{s} = [s(0), s(1), \ldots, s(N_f)]$ is the signal in vector form. Since we have already assumed the prototype filter to be boxcar, $\mathbf{H}$ is the $N_f \times N_f$ identity matrix. By trigonometric identities, we represent $s(k)$ as

$$s(k) = \frac{A}{2}\left[e^{j\frac{2\pi kf}{N_f}} e^{j\delta} + e^{-j\frac{2\pi kf}{N_f}} e^{j\delta}\right] \tag{2.7}$$

Applying the transformation, we recognize that for $f = l$, the inner product of $\mathbf{W}$ and $\mathbf{s}$ will be a constant, but for $f \neq l$, it will be zero by orthogonality. Thus, we have

$$\mathbf{X_s} = Re(\tilde{A}e^{j\delta}) = \tilde{A}\cos(\delta) \tag{2.8}$$

and

$$\mathbf{Y_s} = Im(\tilde{A}e^{j\delta}) = \tilde{A}\sin(\delta) \tag{2.9}$$

for the real and imaginary outputs. The $\tilde{A}$ absorbs the constant due to the DFT gain and polyphase filtering into the original signal amplitude A. Stochastically, $X_s(l)$ and $Y_s(l)$ are characterized by degenerate distributions, as they place all probability at a single point.

We have derived the statistics of a received pulse in the absence of noise. Now, we look at the results of a received signal of noise alone, $r(k) = n(k)$. The signal $n(k)$ is a random i.i.d. sample of size N_f, with each marginal distribution Gaussian. The joint pdf of the random sample is multivariate Gaussian of dimension N_f. A theorem from multivariate distribution theory will be subsequently helpful.

Theorem 2.1 (Linear Transformations of Multivariate Normals) *Let* $\mathbf{X} = [X_1, X_2, \ldots, X_n]$ *be a random sample with a joint pdf that is multivariate normal and is distributed as*

$$\mathbf{X} \sim N_n(\mu, \Sigma) \tag{2.10}$$

where the mean vector is μ *and the covariance matrix is* Σ. *A random variable* $\mathbf{Y}$ *obtained by a linear transformation of the random sample* $\mathbf{X}$, $\mathbf{Y} = \mathbf{LX}$ *is distributed as*

$$\mathbf{Y} \sim N_n(\mathbf{L}\mu, \mathbf{L}\Sigma\mathbf{L}') \tag{2.11}$$

where $'$ *denotes transposition.*

Because the DFT is a linear transformation, we can use the theorem to find the distribution of the real and imaginary outputs due to noise:

$$\mathbf{X_n} = Re(\mathbf{WHn}) \tag{2.12}$$

and

$$\mathbf{Y_n} = Im(\mathbf{WHn}) \tag{2.13}$$

where $\mathbf{W}$ and $\mathbf{H}$ are as before. Thus, the distributions of $\mathbf{X}$ and $\mathbf{Y}$ are given by

$$\mathbf{X_n} \sim N(Re(\mathbf{W})\mu, Re(\mathbf{W})\Sigma Re(\mathbf{W})') \tag{2.14}$$

and

$$\mathbf{Y_n} \sim N(Im(\mathbf{W})\mu, Im(\mathbf{W})\Sigma Im(\mathbf{W})') \tag{2.15}$$

Since the antenna is ac coupled to the ADC, the noise is zero mean, and the covariance matrix is $\sigma^2 I$ by the i.i.d. assumption. Furthermore, since $\mathbf{W}$ is an orthogonal matrix,

$$\mathbf{X_n} \sim N(0, \sigma^2 Re(\mathbf{WW}')) = N(0, \sigma^2) \tag{2.16}$$

and

$$\mathbf{Y_n} \sim N(0, \sigma^2 Im(\mathbf{WW}')) = N(0, \sigma^2). \tag{2.17}$$

To find the distribution of the received signal of the original reduced model form $r(k) = s(k) + n(k)$, we only need to apply the individual results from the signal and the noise alone.

Theorem 2.2 (Convolution Formula) *Let X_1 and X_2 be independent random variables with probability density functions $f_1(x)$ and $f_2(x)$, respectively. Let there be a transformation, $Y_1 = X_1 + X_2$ and $Y_2 = X_2$. Applying a change of variable technique, the joint pdf of Y_1 and Y_2 is*

$$g(y_1, y_2) = f_1(y_1 - y_2) f_2(y_2) \tag{2.18}$$

and the marginal pdf of $Y_1 = X_1 + X_2$ is given by

$$g_1(y_1) = \int_{-\infty}^{\infty} f_1(y_1 - y_2) f_2(y_2)\, \mathrm{d}y_2 \;=\; f_1(x_1) * f_2(x_2). \tag{2.19}$$

Thus, the pdf of the sum of independent random variables is the convolution of the constituent pdf's.

By Theorem 2.2, $\mathbf{X}$ and $\mathbf{Y}$ are distributed as the convolution of the distributions of $s(k)$ and $n(k)$,

$$\mathbf{X_{s+n}} \sim \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2\sigma^2} s(k)^2} * \delta(\tilde{A}\cos(\delta)) = N(\tilde{A}\cos(\delta), \sigma^2) \tag{2.20}$$

and

$$\mathbf{Y_{s+n}} \sim \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2\sigma^2}s(k)^2} * \delta(\tilde{A}\sin(\delta)) = N(\tilde{A}\sin(\delta), \sigma^2) \tag{2.21}$$

so $\mathbf{X}$ and $\mathbf{Y}$ are distributed as normal random variables with equal variance with mean dependent on the initial phase of the signal.

To complete the derivation of the statistics for the reduced model, we must now apply the final step of magnitude conversion. In general, the magnitude of the DFT would be calculated as $\mathbf{Z} = \sqrt{\mathbf{X}^2 + \mathbf{Y}^2}$, but since square roots are very computationally complex in the FPGA, we choose to use the squared magnitude $\mathbf{Z} = \mathbf{X}^2 + \mathbf{Y}^2$ instead. From distribution theory, we know that for squared normal variables, we expect a chi-squared distribution, and in fact, we expect two distributions in our case. For noise alone,

$$\mathbf{Z} \sim \sigma^2 \chi_2^2 \tag{2.22}$$

and for the signal plus noise case,

$$\mathbf{Z} \sim \sigma^2 \chi_2'^2(\lambda) \tag{2.23}$$

where χ_n^2 is the central chi-squared distribution with n degrees of freedom, and $\chi_n'^2(\lambda)$ is the noncentral chi-squared distribution with n degrees of freedom and noncentrality parameter λ. In our case, the noncentrality parameter is

$$\lambda = \left(\frac{\tilde{A}\cos(\delta)}{\sigma}\right)^2 + \left(\frac{\tilde{A}\sin(\delta)}{\sigma}\right)^2 = \frac{\tilde{A}^2}{\sigma^2} \tag{2.24}$$

Thus, we expect chi-squared random variables at the output of our polyphase filter bank.

2.2 General Signal Model

Admittedly, the reduced signal model is not of much use for signals of interest in ELINT applications, but it gives us a basis for extrapolating to a more general signal model that encompasses such variables as pulse modulation, phase modulation, and pulse-length considerations. Of the assumptions of Section 2.1, we retain our assumptions about the noise only, as we will throughout this entire thesis. However, we now allow both phase and pulse modulation of our signal

$$s(k) = A(k)\cos(\omega(k) + \theta(k) + \delta) \tag{2.25}$$

where $A(k)$ is the pulse modulation, and $\theta(k)$ is the phase modulation.

It is excessive to repeat the above derivation of the statistics as we can use the results to make general inferences. The most simple case is that of only pulse envelope modulation, as that would imply that the noncentrality parameter is no longer constant but would be a function of time. In the case of phase modulation, the inner product of the DFT matrix $\mathbf{W}$ with the signal would no longer be constant in one channel, but would be constant in any number of channels, and also a function of time. Both of these results do not change the expected statistics; we retain a chi-squared distribution, but it becomes nonstationary.

2.3 Prototype Filter and Pulse Lengths

In the reduced signal model, we assumed a boxcar filter of length N_f, the DFT length. However, this resulted in no prototype filter at all. In general, we would like a prototype filter that achieves the desired out-of-band rejection characteristics. The filter prototype will usually need some overlap of the DFT to accomplish this, and we assumed above that the pulse length was equal to the filter length. For pulses longer than the filter length, this is not a problem, but for any pulse shorter than the filter length we expect to see a reduction in SNR as the filtering operation spreads the concentrated energy of the pulse to the full length of the filter. In an optimal situation, since the pulse length is unknown, we would run multiple filter banks in parallel, each with a different prototype filter sized to optimally filter a pulse of expected length. However, FPGA implementation prohibits this excessive computation. Thus, we must choose with care the filter length in order to preserve SNR for short pulses, while providing sufficient channel subband characteristics after the transformations.

An important function of the prototype filter is to shape each subband channel, but equally important is the effect of the filtering on the statistics between and within the channels. Our detection methods will rely on statistical inference, or hypothesis testing. In many methods, the statistical independence of the samples is important in determining multivariate distributions. For our purposes, we would like to limit the correlation of the samples as much as possible, in order to declare them independent to the tolerance of our tests. We can use the methods of Section 2.1 and particularly the theorem of transformations of multivariate normals to find the expected correlation of the output variables Z. For the case of a nonboxcar prototype filter of

any length, the distribution of the real and imaginary components in each channel is

$$\mathbf{X} \sim N((\mathbf{WH})A\cos(\delta), \sigma^2(\mathbf{WH})(\mathbf{WH})') \tag{2.26}$$

and

$$\mathbf{Y} \sim N((\mathbf{WH})A\sin(\delta), \sigma^2(\mathbf{WH})(\mathbf{WH})') \tag{2.27}$$

If we work with the covariance matrix $\Sigma = \sigma^2(\mathbf{WHH'W'})$, we can find the correlation coefficients both between and within each channel for any prototype filter and DFT combination.

2.3.1 Intrachannel correlation

By intrachannel correlation, we mean the correlation of the samples between the different channels of the polyphase output, before the magnitude conversion. For example, how is sample k of the eighth channel correlated with that of its neighbor in the seventh or ninth channels? This becomes important in bandwidth estimation later, where we would like statistical independence in order to infer whether there is significant energy in any channel. For the covariance matrix Σ above, statistical independence implies orthogonal transformations. Because the DFT matrix $\mathbf{W}$ is necessarily orthogonal, if we limit the polyphase transformation of the prototype filter $\mathbf{H}$ to be orthogonal we require strict statistical independence between channels. Specifically, we guarantee that the corellation coefficient given by

$$\rho = \frac{Cov(Z_i, Z_j)}{\sigma_i \sigma_j} \tag{2.28}$$

is zero in every other channel for aligned data.

2.3.2 Interchannel correlation

Similarly to intrachannel correlation, by *inter*channel correlation we mean the correlation between different samples in the same channel. For example, how are samples k and $k-1$ of the eighth channel correlated? This becomes important in the pulse detection stage, as we will employ a channelized detector for each channel, independent of the other channels. For this analysis, we must generalize $\mathbf{H}$ so that the linear transformation gives us more than one output sample in each channel for comparison. Here, we will become more specific to this project for ease of notation in terms of specifying a prototype filter and DFT length, but the results can be readily generalized to any length by following a similar methodology. For this project, the

DFT length is 32, and a suitable prototype filter length is 128. Let us define the prototype filter polyphase transformation matrix to be $\mathbf{H}$ of size 32×128, and partition that matrix into four submatrices each of size 32×32 as in

$$\mathbf{H} = \begin{bmatrix} \mathbf{H_1} & \mathbf{H_2} & \mathbf{H_3} & \mathbf{H_4} \end{bmatrix}$$

Now, we desire to yield more than one output sample in each channel. If we construct a new matrix $\mathbf{H}^*$ as in

$$\mathbf{H}^* = \begin{bmatrix} \mathbf{H_1} & \mathbf{H_2} & \mathbf{H_3} & \mathbf{H_4} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{H_1} & \mathbf{H_2} & \mathbf{H_3} & \mathbf{H_4} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{H_1} & \mathbf{H_2} & \mathbf{H_3} & \mathbf{H_4} & \mathbf{0} & \cdots & \mathbf{0} \\ \vdots & & & & & \ddots & & & \vdots \end{bmatrix}$$

and a new DFT matrix $\mathbf{W}^*$ as in

$$\mathbf{W}^* = \begin{bmatrix} \mathbf{W} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{W} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{W} & \mathbf{0} & \cdots & \mathbf{0} \\ \vdots & & & \ddots & & \vdots \end{bmatrix}$$

where $\mathbf{0}$ denotes a matrix of zeros of size 32×32, then we can form a new linear transformation that yields multiple output samples in each channel. This new transformation is given by

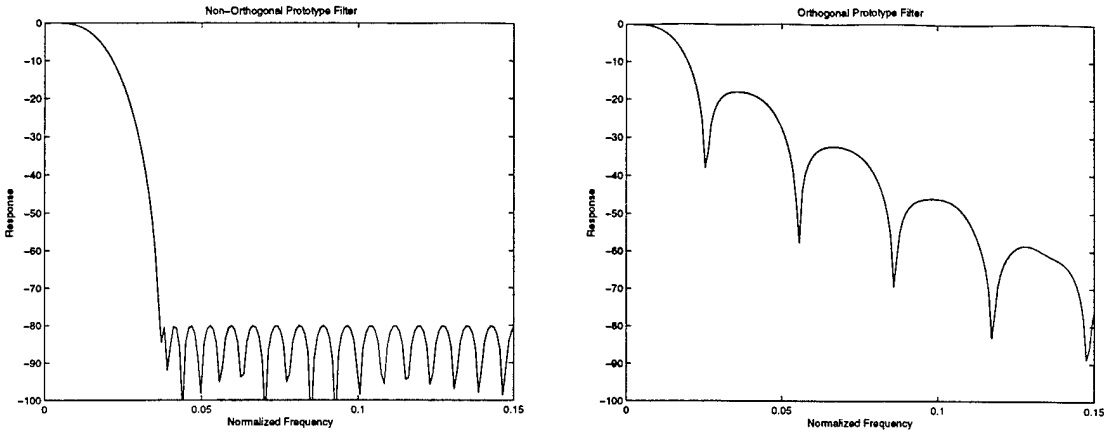
$$\mathbf{X} = Re(\mathbf{W}^*\mathbf{H}^*\mathbf{r}) \tag{2.29}$$

$$\mathbf{Y} = Im(\mathbf{W}^*\mathbf{H}^*\mathbf{r}) \tag{2.30}$$

where $\mathbf{r}$ is the received signal in the presence of noise. We can now compute the covariance matrix $\Sigma^* = \sigma^2(\mathbf{W}^*\mathbf{H}^*\mathbf{H}^{*\prime}\mathbf{W}^{*\prime})$ and use it to find the intrachannel correlation coefficients.

2.3.3 Correlation results

Applying the methods above, we can find the correlation due to any prototype filter. We begin with the filter designed specifically for this project by Joseph Arrowood, which is a length-128 non-orthogonal FIR filter [2]. By orthogonalizing this filter, we can compare the results of the orthogonal and non-orthogonal cases to determine the costs and benefits of each method. In Figure 2.2, we can see that the nonorthogonal version of the prototype filter has



(a) Nonorthogonal Filter

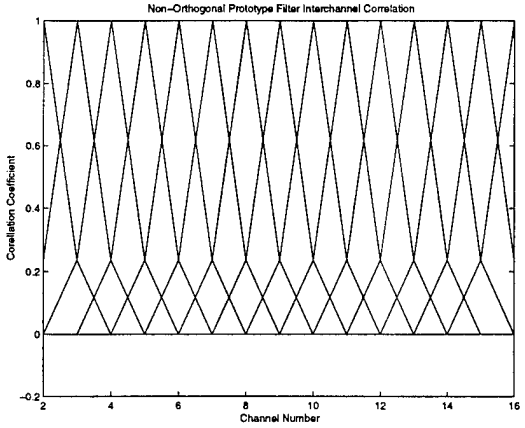
(b) Orthogonal Filter

Figure 2.2 Normalized frequency response of a non-orthogonal filter designed by J. Arrowood, and an orthogonalized version.

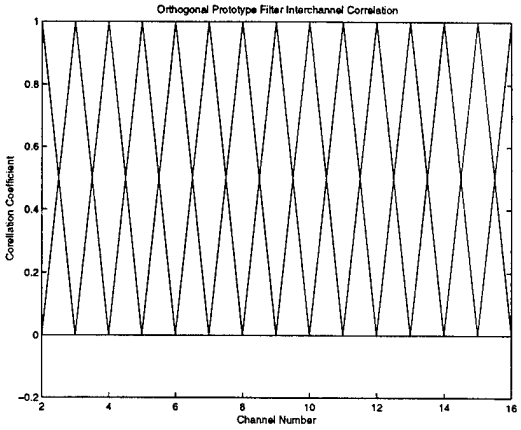
much lower sidebands, resulting in a high out-of-band energy rejection in adjacent channels. This is important in bandwidth measurement applications, as we would like to keep a strict channelization of the energy without excess leakage.

Using these two prototype filters, we can now compare the interchannel and intrachannel correlation coefficients. From the covariance matrix derived in the above section, we can compute the correlation coefficient ρ with Equation (2.28). The correlation coefficient is bounded by $-1 \geq \rho \geq 1$. Statistical independence implies $\rho = 0$, and the closer ρ is to zero, the more "independent" the respective samples are. We put "independent" in quotation marks because for $\rho \neq 0$ the samples are necessarily dependent, but we may treat them as approximately independent for ρ near zero. In Figure 2.3(a),(b), the interchannel correlation coefficient is necessarily zero, as we would expect, in the case of the orthogonal prototype filter, but is about .2 for the non-orthogonal filter. While this is higher than we would like, it remains low enough that we may treat them as approximately independent as a trade-off for high sideband attenuation, which is more important than statistical independence between the channels. In Figure 2.3(c),(d), the intrachannel correlation coefficient is similar for both the non-orthogonal and orthogonal prototypes. This is to be expected due to the nature of the polyphase filtering. For any overlap where the filter length is greater than the DFT length, we have a projection operator that projects the original vector onto a smaller subspace. This data reduction implies that some combination of the data is taking place, resulting in correlation of the output

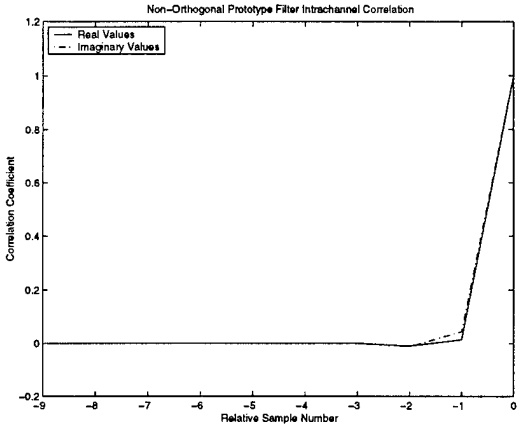
samples. However, we notice that in both cases, the correlation coefficient rapidly approaches zero away from the current sample, and is still small for the adjacent sample. Thus, we can safely approximate statistical independence, especially when $\rho \approx .03$ as in this case for the non-orthogonal filter. From this analysis, the obvious choice is to proceed with a non-orthogonal prototype filter due to its small intrachannel correlation and high sideband attenuation.



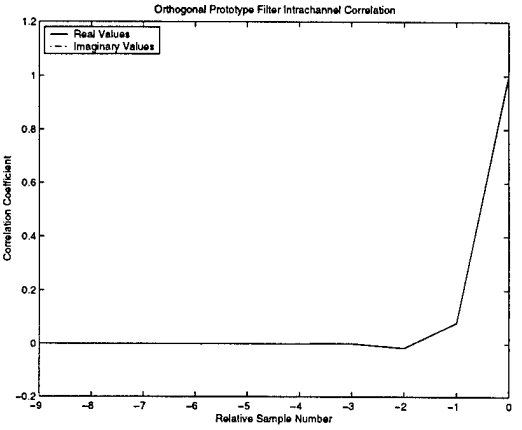
(a) Non-Orthogonal



(b) Orthogonal



(c) Non-Orthogonal



(d) Orthogonal

Figure 2.3 (a),(b) Interchannel correlation coefficients in each of 15 channels for both the imaginary and real parts of each channel. (c),(d) Intrachannel correlation coefficients based on relative sample number, where 0 is the current sample number and previous sample numbers appear as -k for the real and imaginary parts of a representative channel.

CHAPTER 3

FIXED-SAMPLE-SIZE DETECTION METHODS

Traditionally, this type of subband detection of radar pulses has been accomplished with fixed-sample-size (FSS) statistical hypothesis tests. The optimal test is determined by the statistics after the linear transformation that is chosen for the subbanding. Thus, for our case, the optimal test for the output of a polyphase filter bank is necessarily much different from the optimal test for the output of a cosine modulated filter bank, or a wavelet decomposition. As such, we refer to the "optimal" solution as it applies to this particular subband decomposition. In general, there is a truly optimal solution to the radar detection problem utilizing matched filters and multiple independent filter banks for each expected pulse length. However, this is not implementable in today's FPGAs as it is too computationally complex. We will derive and characterize the optimal fixed-sample-size test for the particular case of this subband decomposition. In fact, for this project, the goal of this work was to determine what exactly is the optimal solution given this subband decomposition, which is a standard practice in ELINT applications.

3.1 Optimal Fixed-Sample-Size Detector

Given the statistics present at the output of the polyphase filter bank derived in Chapter 2, we can derive the optimal test for the case of signal plus noise over purely noise. For these detectors, we will assume that the detector works in a single subband, independently of all other channels. For a 32-band decomposition, we would generate 15 usable channels, neglecting the dc and Nyquist bands that are subject to fixed-precision filtering anomalies [2]. Thus, in our detection scheme, there would be 15 FSS detectors running independently in each channel. In general, more complicated methods of alarming on multiple channels at once are possible, and

these will be discussed during bandwidth estimation, as they are more effective and appropriate in that setting.

From Chapter 2, we know that the statistics present at the output of the channelizer are given by

$$Z \sim \sigma^2 \chi_2^2 \qquad \text{(Noise Only Case)} \tag{3.1}$$

and

$$Z \sim \sigma^2 \chi_2'^2(\lambda) \qquad \text{(Signal Plus Noise Case)} \tag{3.2}$$

where we define Z as the output of the polyphase filter bank subject to magnitude conversion. As we saw before, we will treat these Z as statistically independent subject to the considerations of the prototype filter maintaining a small correlation coefficient.

Our first problem in finding an optimal solution is to find a statistic that represents the data well, and separates it into two classes based on its properties. The defining characteristic of our two cases of signal and signal plus noise is the noncentrality parameter λ. This is directly equivalent to a shift in mean, as will be shown later when we derive a consistent estimator of the noncentrality parameter for continuous interference. Thus, our statistic should be a sufficient estimator for λ. Furthermore, we require our tests to be noise-riding. Since we cannot assume stationarity, we must directly calculate the parameters of the chi-squared distribution from the data, in general using maximum likelihood (ML) estimators. We now present formally a method for achieving maximal test power over the class of estimators of a chosen size.

Theorem 3.1 (Neyman-Pearson Theorem) *Let $X_1, X_2, \ldots, X_n$, where n is a fixed positive integer, denote a random sample from a distribution that has PDF $f(x;\theta)$. Then the joint PDF of $X_1, X_2, \ldots, X_n$ is*

$$L(\theta; x_1, x_2, \ldots, x_n) = f(x_1;\theta) f(x_2;\theta) \cdots f(x_n;\theta). \tag{3.3}$$

Let θ' and θ'' be distinct fixed values of θ to that $\Omega = [\theta : \theta = \theta', \theta'']$, and let k be a positive number. Let C be a subset of the sample space such that

(a) $\frac{L(\theta';x_1,x_2,\ldots,x_n)}{L(\theta'';x_1,x_2,\ldots,x_n)} \leq k$ *for each point* $(x_1, x_2, \ldots, x_n) \in C$

(b) $\frac{L(\theta';x_1,x_2,\ldots,x_n)}{L(\theta'';x_1,x_2,\ldots,x_n)} \geq k$ *for each point* $(x_1, x_2, \ldots, x_n) \in C^*$

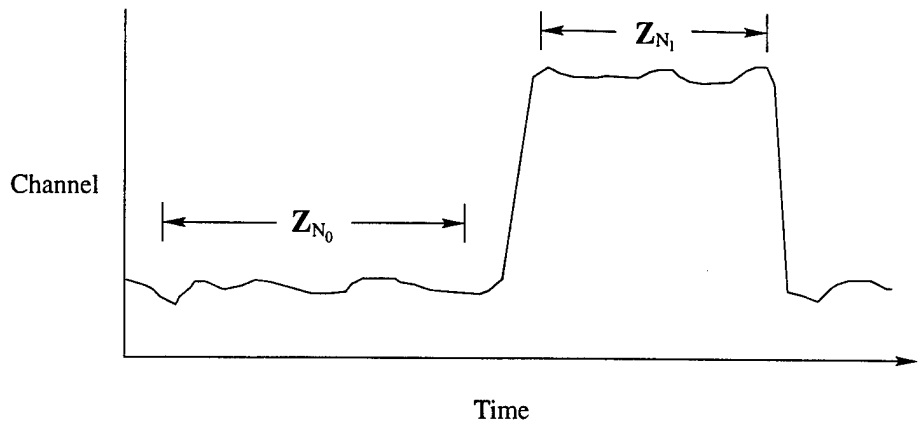


Figure 3.1 Graphical interpretation of the choice of random samples for testing hypothesis H_0 against H_1.

(c) $\alpha = Pr[(X_1, X_2, \ldots, X_n) \in C; H_0]$

Then C is a best critical region of size α for testing the simple hypothesis $H_0 : \theta = \theta'$ against the alternative simple hypothesis $H_1 : \theta = \theta''$ [3].

The Neyman-Pearson theorem states that we can form our test of a simple hypothesis on one parameter of a given distribution if we can construct a likelihood ratio $L(\theta; x_1, x_2, \ldots, x_n)$, based on a sufficient statistic of a set of i.i.d. random variables. Furthermore, if we can find a constant k such that the size of the test is α, and the constant divides the hypothesis space into mutually exclusive regions, C and C^*, then the test is a best test of the hypothesis if the size of C^* is as great as with any other test. Intuitively, we may think of the size of the critical region C as the false alarm probability α, and the size of the complementary critical region C^* as the probability of detection β, or the power of the test. Thus, Neyman-Pearson gives us a way of maximizing the detection probability for a given false alarm rate (FAR).

Suppose now that we base our statistic on a sum of i.i.d. random variables. Ideally, we can compare two random samples of the channelized data and look for a significant difference between their noncentrality parameters. As in Figure 3.1, it makes sense to take independent observations of data; one will form the basis of the likelihood ratio for the noise-only hypothesis, and a second random sampling that will be tested against the first observation. We define mathematically two observation vectors: $\mathbf{Z_{N_0}} = [Z_0, Z_1, \ldots, Z_{N_0}]$ of size N_0 and a second observation $\mathbf{Z_{N_1}} = [Z_{N_0+r}, Z_{N_0+r+1}, \ldots, Z_{N_0+r+N_1}]$ of size N_1 where r is a constant chosen to

separate the observations by the average rise time allowed by the input bandwidth. By the Neyman-Pearson theorem, we construct the likelihood ratio as the joint pdf of the observations $\mathbf{Z_{N_1}}$ over $\mathbf{Z_{N_0}}$ which are each i.i.d. distributed chi-square variables. Thus, the null hypothesis H_0 is that the noncentrality parameter λ_0 estimated from $\mathbf{Z_{N_0}}$ is equal to the noncentrality parameter λ_1 estimated from the observations $\mathbf{Z_{N_1}}$. Thus, the statistic

$$\mathcal{F} = \frac{\frac{1}{N_1} \sum_{i=N_0+r}^{N_0+r+N_1} Z_i}{\frac{1}{N_0} \sum_{j=0}^{N_0} Z_j} \tag{3.4}$$

represents the likelihood ratio

$$\mathcal{F} = \frac{L(\lambda_1; Z_{N_0+r}, Z_{N_0+r+1}, \ldots, Z_{N_0+r+N_1})}{L(\lambda_0; Z_0, Z_1, \ldots, Z_{N_0})} \tag{3.5}$$

It is worth noting that this method is not optimal under all circumstances. In Chapter 2, we found the correlation coefficient to be small due to the filtering operations, but this is under the assumption that the input is i.i.d. Under the H_0 hypothesis, the observations $\mathbf{Z_{N_0}}$ and $\mathbf{Z_{N_1}}$ are independent, and the likelihood ratio constructed from the sufficient statistic of the sum of the observed data is optimal. However, under the H_1 hypothesis, there may be dependence due to the received signal because it is not necessarily a process which is random or independent. In these circumstances, improved performance can be obtained by coherently summing the energy and then squaring the result. However, in cases such as the general signal model of Chapter 2 in which we allow a center frequency of the received pulse to be off the center of any band of the decomposition and we allow phase modulation, or possibly phase drift due to Doppler effects, this sufficient statistic is near optimal since we can no longer coherently sum the energy across DFT blocks, and must resort to incoherent averaging to achieve the best performance.

3.2 Detection without Interference

We will now show how the general results of the likelihood ratio above can be categorized into special cases that yield significant results. We have formed our likelihood ratio from the Neyman-Pearson theorem, but to achieve a best test we must find some way to calculate the constant k such that we produce a critical region of size α, while maximizing the power of our test β.

Consider the special case of $\lambda_0 = 0$. This could be considered our standard operating mode under the assumptions made in Chapter 2 as we assumed that the noise was zero mean. Thus, under the null hypothesis, we are testing $\lambda_1 = \lambda_0 = 0$, and wish to reject this hypothesis only at a significance level α. Under these assumptions with the null hypothesis valid, our likelihood ratio is distributed as

$$\mathcal{F} \sim \frac{\chi^2_{2N_1}/N_1}{\chi^2_{2N_0}/N_0} = F_{2N_1,2N_2} \tag{3.6}$$

where F_{ν_1,ν_2} is the F distribution with ν_1 and ν_2 degrees of freedom [4]. The F pdf is given by

$$f_F(x) = \frac{\Gamma\left(\frac{\nu_1+\nu_2}{2}\right)}{\Gamma\left(\frac{\nu_1}{2}\right)\Gamma\left(\frac{\nu_2}{2}\right)}\left(\frac{\nu_1}{\nu_2}\right)^{\nu_1/2}\frac{x^{\frac{\nu_1-2}{2}}}{\left(1+\frac{\nu_1}{\nu_2}x\right)^{\frac{\nu_1+\nu_2}{2}}} \qquad \text{for } x > 0 \tag{3.7}$$

Thus, since the distribution of the likelihood ratio is known, we can use it to calculate the constant

$$k = F_{2N_1,2N_2;\alpha} \tag{3.8}$$

where $F_{2N_1,2N_2;\alpha}$ is calculated from the inverse F cdf such that condition (c) of the Neyman-Pearson theorem is valid. These F values are easily found in standard tabulations and are also a part of many mathematics packages including MATLAB, S-plus, SAS, and in C or Fortran routines. A test of this form is commonly called an F-test and forms the basis of much statistical theory.

3.3 Optimal Properties of the F-test

This section is primarily interested in explaining in a mathematical sense why the F-test has desirable properties for FSS detection given the statistics we expect at the output of the polyphase filter bank. Let us denote the hypothesis H_0 such that $\lambda_1 = 0$ under the assumption that there is no interference and $\lambda_0 = 0$. We denote the n-dimensional sample space V_n as the vector of observations $\mathbf{Z} = [Z_1, Z_2, \ldots, Z_n]$. The choice of a test of H_0 is equivalent to the choice of a Borel set C in V_n, rejecting H_0 if and only if the observed point $\mathbf{Z}$ falls into C^*.[5] The power of the test is given by

$$\beta(\lambda_1, C^*) = \int_{C^*} p_\lambda(\mathbf{Z})\,\mathrm{d}\mathbf{Z} \tag{3.9}$$

which is a function of the noncentrality parameter and the critical region C^* only.

Definition 3.1 (Uniformly Most Powerful Region) *C^* is a uniformly most powerful (UMP) critical region of a given class $\mathcal{A}$ if $C^* \in \mathcal{A}$, and if, for any $C^{*'} \in \mathcal{A}$, $\beta(\lambda_1, C^*) \geq \beta(\lambda_1, C^{*'})$ [3].*

We know that the power of the test depends on λ_1 only through the intermediary of the means of the Gaussian variables $\mathbf{X}$ and $\mathbf{Y}$ of Chapter 2 which are calculated from the amplitude of the pulse. The multivariate noncentrality parameter is $\sum_{k=1}^{n} \tilde{A}_k^2 = \sigma^2 \lambda'^2$, where λ' is any nonnegative constant. This is equivalent to saying that the power is constant on surfaces

$$\mathcal{B}_{\lambda'} = \left[\lambda_1 \middle| \sum_{k=1}^{n} \tilde{A}_k^2 = \sigma^2 \lambda'^2 \right] \tag{3.10}$$

The surface $\mathcal{B}_{\lambda'}$ is a hypercylinder whose base is a spherical cone of one nappe [5].

Theorem 3.2 (Uniformly Most Powerful F-test) *Among all critical regions C of size α with the property that the power depends only on λ_1 through the intermediary of $\sigma^2 \lambda'^2$, C is UMP [5].*

We now introduce a space of n-dimensional spheres

$$\mathcal{S}(\tilde{A}'_1, \tilde{A}'_2, \ldots, \tilde{A}'_n, \lambda') = \left[\lambda_1 | \sum_{k=1}^{n} \tilde{A}_k^2 = \sigma^2 \lambda'^2; \tilde{A}'_1 = \tilde{A}_1, \tilde{A}'_2 = \tilde{A}_2, \ldots, \tilde{A}'_n = \tilde{A}_n \right] \tag{3.11}$$

where the constants $\tilde{A}'_1, \ldots, \tilde{A}'_n, \lambda'$ are chosen with only the restriction that $\lambda' \geq 0$ and in our case $\tilde{A}'_i \geq 0$. We define the average power over the sphere $\mathcal{S}$ under the critical region C to be the integral of the power $\beta(\lambda_1, C^*)$ over $\mathcal{S}$. The cylinder $\mathcal{B}$ can be expressed as a union of these spheres [5].

Theorem 3.3 (Average Power Maximization) *In the class of all similar regions of fixed size α for testing H_0, C^* maximizes the average power on every sphere $\mathcal{S}(\tilde{A}'_1, \tilde{A}'_2, \ldots, \tilde{A}'_n, \lambda')$ [5].*

The F-test is thus optimal at maximizing the average power of any test for all alternatives $\lambda_1 > \lambda_0$, resulting in a uniformly most powerful test when we are interested "uniformly" in all alternatives, represented by a uniform weighting of the spheres. In a case where we are interested possibly in one alternative over other possible alternatives, the F-test may be suboptimal and other tests may be superior.

Thus, we find that the F-test has desirable properties for our application. It maximizes the average probability of detection for all SNRs, when we are interested uniformly in all

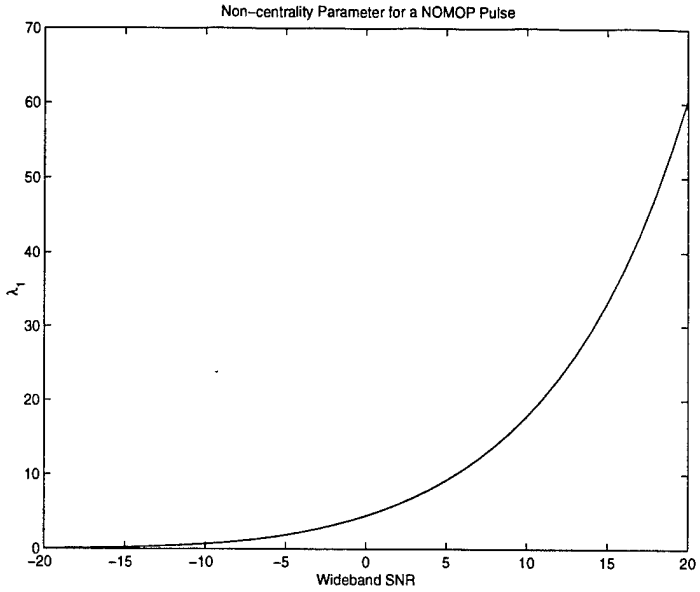


Figure 3.2 Noncentrality parameter λ_1 as a function of wideband SNR for a NOMOP pulse.

alternatives, implying that we are interested in detecting a 3-dB pulse as much as say, a 20-dB pulse. In the event that there is a nonuniform preference for detection, the F-test may be suboptimal.

3.4 Performance of the F-test

The performance of the F-test can be characterized by the power with which it rejects the hypothesis of noise only, H_0, or equivalently by its probability of detection over the range of pulse SNRs. To calculate the power function, we must find the distribution of the statistic $\mathcal{F}$ under the alternative hypothesis H_1 signal present. In this case, we know from our derivation of Chapter 2 that in the presence of signal, the joint pdf of the observation vector $\mathbf{Z}_{\mathbf{N}_1}$ will be noncentral chi-squared with some noncentrality parameter $\lambda_1 > 0$. In general, we found this noncentrality parameter to be a function of the prototype filter H, so we must calculate it based upon its particular gain and other characteristics. Using the non-orthogonal prototype filter discussed in Chapter 2, we simulate the range of SNRs relevant for this project, -20-dB to 20-dB, and find the mean noncentrality parameter for a NOMOP pulse. In Figure 3.2, we see the noncentrality parameter as a function of SNR, which ranges from .07 for a -20-dB pulse to 60.5 for a 20-dB pulse. In this case, we have assumed a NOMOP pulse, but the results

generalize easily to a case where the energy is distributed among more than one channel due to modulation. In these cases, the apparent SNR in each channel may be less than the wideband SNR, but the noncentrality parameter will reflect the apparent channel SNR, and the power associated with the F-test at that SNR will apply similarly.

To find the power of the test, we note that the distribution of the test statistic $\mathcal{F}$ is

$$\mathcal{F} \sim \frac{\chi'^2_{2N_1}(N_1\lambda_1)/N_1}{\chi^2_{2N_0}/N_0} = F'_{2N_1,2N_2}(N_1\lambda_1) \tag{3.12}$$

where $F'_{\nu_1,\nu_2}(\lambda)$ is the noncentral F distribution with ν_1 and ν_2 degrees of freedom with noncentrality parameter λ. The noncentrality parameter as well as the degrees of freedom are heavily influenced by the lengths of the observation intervals for $\mathbf{Z_{N_0}}$ and $\mathbf{Z_{N_1}}$ due to a property of addition of noncentral chi-squared variables.

Property 3.1 (Addition of χ'^2 Random Variables) *Let X_1 and X_2 be independent random variables distributed as $\chi'^2_m(\delta_1)$ and $\chi'^2_n(\delta_2)$, respectively. Then the sum $Y = X_1 + X_2$ is distributed as [4]*

$$Y \sim \chi'^2_{m+n}(\delta_1 + \delta_2) \tag{3.13}$$

For the hypothesis H_1, signal present, this is the distribution of the statistic $\mathcal{F}$. Thus, the power of the test, or the probability of rejecting the hypothesis of signal absent when it is false, is equal to the integral

$$\beta(\lambda_1, C^*) = \int_k^\infty p_{F'}(x)\,\mathrm{d}x \tag{3.14}$$

which is the area under the noncentral F pdf over the critical region C^*. This critical region is the complement of the space C which is the region for accepting H_0 when it is true. In general, the power of the test is dependent not only on the noncentrality parameter but also the critical region specified by the size of the test α, as well as the lengths of the observation samples N_0 and N_1. In Figure 3.3 we calculate the power of the test for different sizes ranging from an α of 10^{-2} to 10^{-5} and observation intervals $N_1 = 4$ and $N_0 = 64$.

3.5 Observation Interval Lengths

From the noncentral F distribution under the hypothesis H_1 we see that both the degrees of freedom as well as the noncentrality parameter influence the distribution of the statistic.

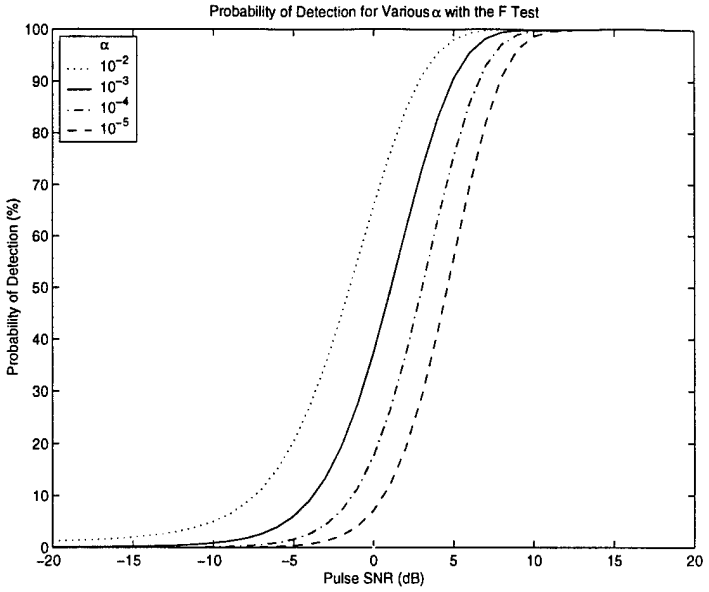


Figure 3.3 Probability of detection (power) of the F-test for various test sizes α and observation lengths $N_1 = 4$ and $N_0 = 64$.

Intuitively, we expect that longer window lengths improve detection probability as they lower the variance in the statistic $\mathcal{F}$. In Figure 3.4, we see that as the window length N_1 gets longer, the probability of detection improves at low SNR.

However, care must be taken not to overestimate the observation interval length N_1, as this will result in a reduction in the apparent SNR of a pulse shorter than N_1. In such cases, the apparent noncentrality parameter $\tilde{\lambda}$ will be

$$\tilde{\lambda} = \frac{L}{N_1}\lambda_1 \tag{3.15}$$

where L is the true length of the pulse. An example is given in Figure 3.5, where a 10-dB, 1.28 μs pulse is overestimated by up to 10 times. Notice that even overestimating the window length by a factor of 2 appears as less than half the true SNR to the detector. In all cases, N_1 must be shorter than the observation interval $\mathbf{Z_{N_0}}$ in order to keep the number of degrees of freedom sufficient in the denominator of the F-test to achieve reliable results.

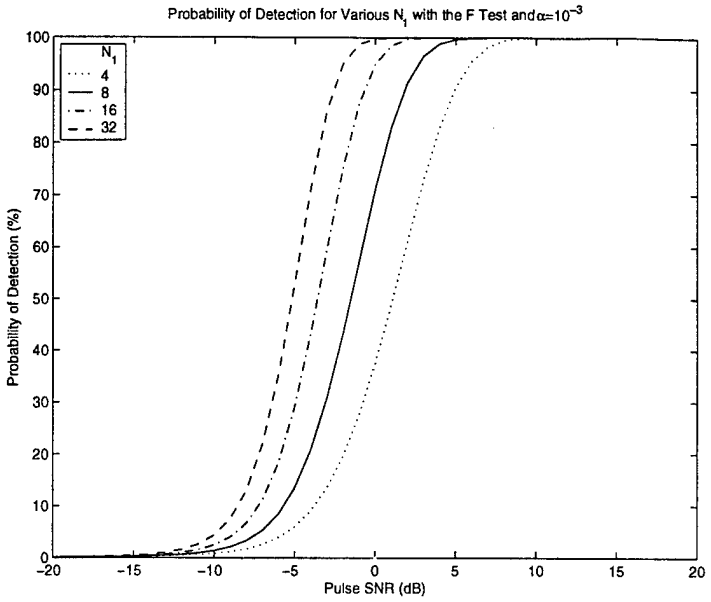


Figure 3.4 Probability of detection (power) for various length of the $\mathbf{Z}_{\mathbf{N}_1}$ observation interval with $N_0 = 64$.

3.6 Detection in the Presence of Interference

In many practical instances the detector may have to perform in the presence of interference such as that from continuous wave (CW) carriers due to civilian communications or other radars that are not interesting for ELINT purposes and are preferably ignored. In general, this interference takes the form of the pulses we would like to detect, such as in the general signal model of Chapter 2, but are of longer duration. In this case, to calculate the constant k and the power of the associated test, we must again find the distribution of the statistic $\mathcal{F}$. In this case, $\mathcal{F}$ is the ratio of two noncentral chi-squared variables and is distributed as

$$\mathcal{F} \sim \frac{\chi'^2_{2N_1}(N_1\lambda_1)/N_1}{\chi'^2_{2N_0}(N_0\lambda_0)/N_0} = F''_{2N_1,2N_0}(\lambda_1,\lambda_0) \tag{3.16}$$

where $F''_{\nu_1,\nu_2}(\lambda_1,\lambda_2)$ is the doubly noncentral F distribution with degrees of freedom ν_1 and ν_2 with noncentrality parameters λ_1 and λ_2. Essentially, the F tests previously were generalizations of this distribution with one or both noncentrality parameters equal to zero. The doubly

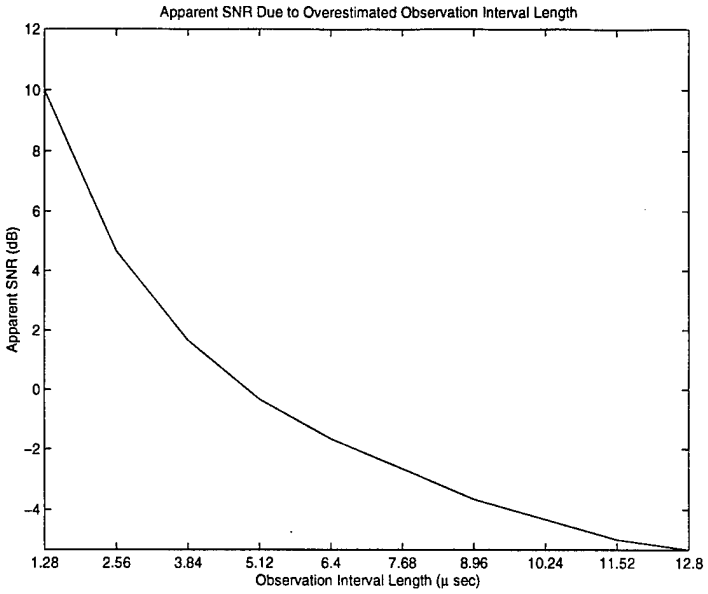


Figure 3.5 Apparent SNR of a 10-dB pulse 1.28 μs long caused by overestimating the observation interval length up to 10 times.

noncentral F PDF is given by

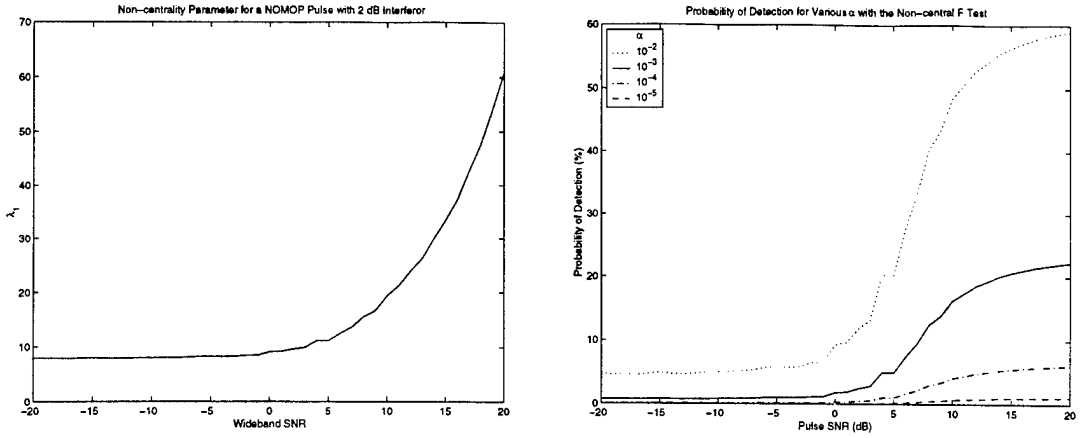
$$p_{F''}(x) = e^{-\frac{1}{2}(\lambda_1+\lambda_2)} \sum_{r=0}^{\infty}\sum_{s=0}^{\infty} E(\lambda_1, r)E(\lambda_2, s) \frac{\Gamma\left(\frac{\nu_1+\nu_2}{2}+r+s\right)}{\Gamma\left(\frac{\nu_1}{2}+r\right)\Gamma\left(\frac{\nu_2}{2}+2\right)} \left(\frac{\nu_1}{\nu_2}\right)^{\frac{m}{2}+r} \frac{x^{\frac{\nu_1}{2}+r-1}}{\left(1+\frac{\nu_1}{\nu_2}x\right)^{\frac{\nu_1+\nu_2}{2}+r+s}} \tag{3.17}$$

where

$$E(\lambda, m) = \frac{\left(\frac{\lambda}{2}\right)^m}{\Gamma(m+1)} \tag{3.18}$$

valid for $x > 0$ and $\lambda_1 \geq \lambda_0$ [6],[7]. In this case, under both hypotheses, the statistic is distributed as F'', with $\lambda_1 = \lambda_0$ under H_0, and $\lambda_1 > \lambda_0$ under H_1. To find the constant k in the Neyman-Pearson theorem, we note that a property of F'' is that for $\lambda_1 = \lambda_0$, as under the null hypothesis, F'' reduces to the noncentral F distribution F', allowing easy calculation of k via mathematics software, as in the case of the F-test. However, for the general case of an unknown SNR interferer, the noncentrality parameter λ_0 must be calculated from the observations of $\mathbf{Z_{N_0}}$. Given the moment generating function of $\chi'^2_\nu(\lambda)$, $M(t;\nu,\lambda)$ it can be shown that the sample mean is

$$\bar{X} = \left.\frac{\partial M(t;\nu,\lambda)}{\partial t}\right|_{t=0} = \nu + \lambda \tag{3.19}$$



(a) Mean Noncentrality Parameter

(b) Power

Figure 3.6 Mean noncentrality parameter and probability of detection (power) of the noncentral F-test for various test sizes α and observation lengths $N_1 = 4$ and $N_0 = 64$ in the presence of a 2-dB CW interferer.

giving a consistent estimator for λ_0 as $\widehat{\lambda_0} = \bar{X} - 2$. This test is commonly referred to as the noncentral F-test. As in the case of no interference, the power of the test is calculated as

$$\beta(\lambda_1, C^*) = \int_k^\infty p_{F''}(x)\,\mathrm{d}x \tag{3.20}$$

using numerical integration. In practice, the numerical simulation of Equation (3.17) is sensitive to both the magnitude of λ_1 as well as to the upper limit of summation over r and s. For large λ_1 it will not converge to a meaningful distribution, and an appropriate upper bound for r and s is approximately 30 to 50.

The power calculation is complicated through the possible constructive or destructive interference of the interferer with the desired pulse. Under the assumption that the initial phases of both the interferer and the pulse are independently uniform on the interval $[0, 2\pi)$, it is possible to achieve anywhere from totally constructive to totally destructive interference. As such, unlike the previous cases, even a NOMOP pulse will have a noncentrality parameter $\lambda_1(t)$ that is a function of time. As an example, simulation was used to find the mean noncentrality parameter of a 3 μs pulse in the presence of a 2-dB interferer when both the interferer and desired pulse have unknown initial phase. In Figure 3.6, the power of the noncentral F test is calculated using simulation. The values are only approximate and are subject to the variance of

numerical integration and the fluctuation of the estimated mean of λ_1. However, it is instructive in that it shows how detection is severely influenced by interferers, as desensitization of the receiver results. Even though the noncentrality parameter for a 20-dB desired pulse is large, the noncentral F-test performs poorly because of the nonlinear relationship of the noncentrality parameter to the shape of the distribution. Even at large SNR, we cannot achieve sufficient separation between the F'' PDF under the H_1 hypothesis and the F' PDF of the H_0 hypothesis.

3.7 FPGA Implementation

For the optimal FSS tests presented in this chapter, FPGA implementation is relatively efficient. In the case of the F-test, a single value of k can be stored and is applicable to all tests, but in the case of an interferer, there will be a critical value k for every SNR possible for the interferer. This is not a problem for FPGA implementation as the noncentrality parameter is calculated from the sample mean and the corresponding critical value is fetched from a lookup table. For the FSS detectors presented here, the operation count yields $N_0 + N_1$ additions, two multiplications, and two divisions. The divisions are most easily accomplished via bit shifting when N_0 and N_1 are of the form 2^n.

CHAPTER 4

SEQUENTIAL DETECTION METHODS

In Chapter 3 on fixed-sample-size methods, the underlying assumption was that in order to implement an FSS test, we must know the fixed-sample-size that fits the data. In ELINT applications, this is precisely the value that has the least a priori knowledge attached to it. We make assumptions regarding the bandwidth, center frequency, and modulations we desire to be captured, but the length of pulses is less well known. We only assume that the shortest pulse will be one sample long, and a longest pulse designation is only for the convenience of excluding CW carriers from our data collection. As such, there is no one "best" fixed-sample-size with which to accomplish detection. FSS methods perform poorly when the observation interval is wrong, by either reducing the effective SNR of received pulses, or performing suboptimally when the window length is too short.

A sequential test attempts to rectify this situation. In this method, after each observation one of three decisions is made: accept hypothesis H_0, reject hypothesis H_0, or continue the test with another observation. As such, the length of the test n is a random variable, dependent on the data. For ELINT applications, this is particularly of interest as the testing procedure adapts itself to the data being received. Intuitively, we would assume that for a high SNR pulse, we may not need as many observations to reach a conclusion, but for low SNR pulses, provided that the pulse is of sufficient length, we could improve performance by allowing the test procedure more samples upon which to make its decision.

The difference between FSS methods and sequential methods is seen most intuitively by noting the way in which each partitions the parameter space. In the last chapter, we denoted the critical region C as the area under which the acceptance of H_0 was strongly preferred. In general, let us define an n-dimensional parameter space by R_n. In FSS testing, according to

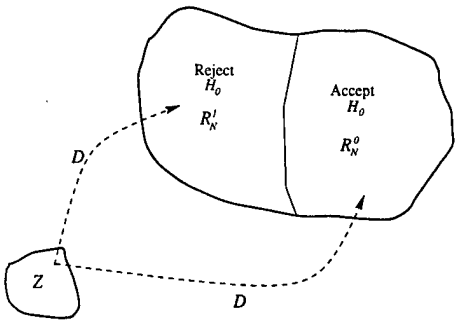


Figure 4.1 FSS partitioning of the N-dimensional parameter space into two mutually exclusive regions corresponding to the size and power of the test.

the Neyman-Pearson theorem, for a fixed size α, we maximize the power of the test β. Thus, we have split the parameter space into two mutually exclusive regions. For a fixed test of length N, R_N^0 is of size α and is the zone of preference for acceptance of H_0, and R_N^1 is the zone of preference for rejection of H_0. In general, it is a decision rule D which partitions the parameter space in this way by means of a sufficient statistic and critical values associated with the boundaries of each space. Graphically, we represent this process by a transformation D which takes the observation space Z and projects it into our divided parameter space to make a decision, as in Figure 4.1. Now, let us consider the partitioning of the parameter space under sequential procedures. Here, the space is divided into three mutually exclusive regions. Since the test length n is a random variable, the parameter space changes dimension with each successive sample, as do the shapes and sizes of the critical regions. We define another decision rule D^* which takes the sample space and projects it into one of three regions. The first region is the zone of acceptance of H_0 and is denoted by R_m^0, which can be any size, as in the Neyman-Pearson formulation. Similarly, we have a zone of preference for the rejection of hypothesis H_0, R_m^1 which, unlike FSS methods, can be made any size, since we now have another region which can cover all other possibilities. This region R_m is the region of indifference, in which it is preferable neither to reject nor accept the hypothesis, but rather to continue observation, as we are yet unsure, based on the observations, what the correct course of action is. Graphically, we show the partitioning in Figure 4.2. The intuitive benefit of sequential testing is the ability to make the critical regions of any size, while leaving the remaining portion of the parameter space to a zone of indifference. In this paper, we will only consider tests which have probability

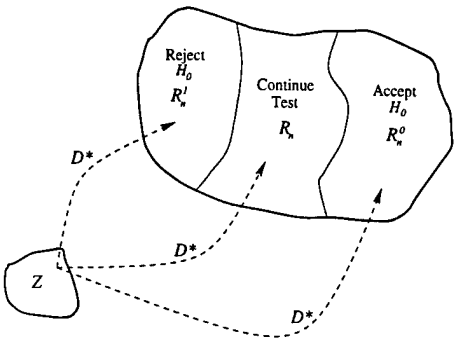


Figure 4.2 Sequential test partitioning of the parameter space into three mutually exclusive regions, the zone for preferable acceptance of H_0, the zone for preferable rejection of H_0, and a zone of indifference where more observations are taken until absorption.

1 of terminating. In other words, we may think of sequential methods as a way of adaptively computing the test size based on the data, which is a desirable property for ELINT.

4.1 Sequential Probability Ratio Test

So far, we have not said how these tests are constructed or what their properties are with respect to test length and SNR detection performance. In pulsed radar detection, the test length is important to the performance of our receiver. For example, if optimal sequential detection required 40 samples to detect on average, but on average we only expected to receive 20 samples from each pulse, then we would conclude that despite its benefits, sequential testing does not perform to the required standard. Sequential testing methods are primarily the result of work by A. Wald [8] who is considered the originator of much of the statistical theory regarding sequential testing. However, the idea of sequential detection has been developed by several others including Page [9] and Lorden [10]. The basic theory behind sequential testing will be drawn from their work, and we will leave out some of the finer mathematical points not of particular interest to us, but will point out where more information regarding certain peculiarities can be found in these works.

Let $L_n(\theta; Z_1, Z_2, \ldots, Z_n)$ be the likelihood of the observed samples $\mathbf{Z_n} = [Z_1, Z_2, \ldots, Z_n]$ coming from a distribution with parameter θ. Then we could construct a likelihood ratio for testing the probability of a sequence of observed points to determine whether they were more

likely to come from one of two possible distributions, as in

$$l_n(\theta; \mathbf{Z_n}) = \frac{L_n(\theta; \mathbf{Z_n})}{L_n(0; \mathbf{Z_n})} \tag{4.1}$$

where the null hypothesis distribution parameter is assumed to be zero in the signal-absent case. This is not in general necessary, as signal absent may be a distribution whose parameter is not zero, but the results are general to any distribution with any parameter, provided that it is a complete family of distributions. The length of the sample set $\mathbf{Z_n}$ is a random variable as in all sequential procedures. We now define our decision rule D^* such that we continue taking observations whenever

$$\mathcal{B} < l_n(\theta; \mathbf{Z_n}) < \mathcal{A} \tag{4.2}$$

for two positive constants $\mathcal{A}$ and $\mathcal{B}$ provided $\mathcal{B} < \mathcal{A}$. If

$$l_n(\theta; \mathbf{Z_n}) \leq \mathcal{B} \tag{4.3}$$

the test is terminated with the acceptance of H_0, and if

$$l_n(\theta; \mathbf{Z_n}) \geq \mathcal{A} \tag{4.4}$$

the test is terminated with the rejection of H_0. In this manner, we divide the parameter space into three mutually exclusive regions with regard to our preference for acceptance and rejection, as well as the zone of indifference. The likelihood ratio $l_n(\theta; \mathbf{Z_n})$ can be expressed as the joint pdf of the samples of $\mathbf{Z_n}$ as in

$$l_n(\theta; \mathbf{Z_n}) = \frac{f(Z_1, \theta) f(Z_2, \theta) \cdots f(Z_n, \theta)}{f(Z_1, 0) f(Z_2, 0) \cdots f(Z_n, 0)} \tag{4.5}$$

where every sample is i.i.d. with marginal pdf $f(Z_i, \theta)$. Suppose that we were to take the logarithm of the likelihood ratio to produce

$$\log l_n(\theta; \mathbf{Z_n}) = \log \frac{f(Z_1, \theta)}{f(Z_1, 0)} + \log \frac{f(Z_2, \theta)}{f(Z_2, 0)} + \cdots + \log \frac{f(Z_n, \theta)}{f(Z_n, 0)} \tag{4.6}$$

We will denote the ith term of this sum as

$$z_i = \log \frac{f(Z_i, \theta)}{f(Z_i, 0)} \tag{4.7}$$

We can reformulate the decision rule D^* by replacing the constants $\mathcal{A}$ and $\mathcal{B}$ with $\log \mathcal{A}$ and $\log \mathcal{B}$ to achieve the same partitioning as above. If

$$\log \mathcal{B} < z_1 + z_2 + \cdots + z_n < \log \mathcal{A} \tag{4.8}$$

then continue the test with sample statistic z_{n+1}, but if

$$z_1 + z_2 + \cdots + z_n \leq \log \mathcal{B} \tag{4.9}$$

the null hypothesis is accepted, and if

$$z_1 + z_2 + \cdots + z_n \geq \log \mathcal{A} \tag{4.10}$$

then the null hypothesis is rejected. This test is referred to as the sequential probability ratio test (SPRT) as developed by Wald [8].

4.2 Determining $\mathcal{A}$ and $\mathcal{B}$

As in constructing all statistical tests, we must determine the critical constants that appropriately size the parameter spaces R_n^0, R_n^1, and R_n. Consider first the rejection of the null hypothesis when

$$l_n(\theta, \mathbf{Z_n}) \geq \mathcal{A} \tag{4.11}$$

Essentially we are saying that the joint probability measure $f(Z_1, \theta)f(Z_2, \theta), \cdots, f(Z_n, \theta)$ is at least $\mathcal{A}$ times as large as the probability measure $f(Z_1, 0)f(Z_2, 0), \cdots, f(Z_n, 0)$. We defined earlier the power of the test as β when H_1 is true, which is the probability of rejecting the null hypothesis when it is false. However, there is associated with that the size of the test α which is the probability of rejecting the null hypothesis when it is true, or a false alarm. The joint probability measures above therefore state that for the hypothesis to be rejected, the power must be $\mathcal{A}$ times the false alarm probability. Thus, we generate the following inequality:

$$\mathcal{A} \leq \frac{\beta}{\alpha} \tag{4.12}$$

Similarly, for a case where the null hypothesis is accepted,

$$l_n(\theta, \mathbf{Z_n}) \leq \mathcal{B} \tag{4.13}$$

The probability measure under the null pdf must be $\frac{1}{\mathcal{B}}$ times as large as that under the alternative pdf. Associating the power and size of the test with this inequality we get

$$\mathcal{B} \geq \frac{1-\beta}{1-\alpha} \tag{4.14}$$

These inequalities yield the bounds of the constants, but we are interested more in determining their values for practical application.[†] For the above derivation we have found inequalities, but what is the outcome if we were to set the constants $\mathcal{A}$ and $\mathcal{B}$ to equalities in practice. It is sufficient to say in this instance that the determination of the exact power and size based on any two constants is quite tedious. Wald provides a method for this type of calculation, but also finds that for the strict equality, the size and power of the test do not change by an appreciable amount when the test is untruncated. The problem in determining the exact error bounds lies in what is called the "excess over boundaries" problem. In the above derivation, we assumed that the for the size and power of the test to have exact values α and β, we achieved the bound $\mathcal{A}$ or $\mathcal{B}$ with strict equality. However, this might require a noninteger number of observations. When we are restricted to integer observations, as we always are, we may exceed the boundaries by a small amount, either decreasing the size or power of the test. Strict equality of the above constants to their derived values does yield one bound that is exact. That is, for a choice of

$$\mathcal{A} = \frac{\beta}{\alpha} \tag{4.15}$$

and

$$\mathcal{B} = \frac{1-\beta}{1-\alpha} \tag{4.16}$$

either the size or power of the test will be exact, and the other will be decreased by a small amount due to the excess over the boundaries.

4.3 Operating Characteristic Function

In FSS methods, the probability of detection is a function of the size of the test and the limiting distribution of the likelihood statistic $\mathcal{F}$. This is easy to find and simulate via mathematical software because the length of the test is fixed. For sequential methods, however, in which the observation interval is itself a random variable, the distribution of the statistic is a function of the interval and the decision function. As such, asymptotic approximations to the actual performance of the test are the methods by which the tests are characterized. Suppose

[†]The reader familiar with Wald's work may notice the difference between Wald's bounds of $\mathcal{A} = \frac{1-\beta}{\alpha}$ and $\mathcal{B} = \frac{\beta}{1-\alpha}$ from those presented here. There is no difference, however, except in notation. Earlier, we presented the power of the test as β, but in Wald's derivation the power of the test is $1-\beta$ with β being the probability of a Type II error. Thus, we have used β here to denote power to be consistent with earlier sections, but the bounds remain the same regardless of the different formulation.

we have a function $\mathcal{L}(\theta)$ which is the conditional probability of accepting hypothesis H_0 at the parameter point θ. Then for no signal, or $\theta = 0$, $\mathcal{L}(\theta) = 1 - \alpha$ and is otherwise a function of the parameter θ, $\mathcal{L}(\theta) = 1 - \beta(\theta)$. Thus, the probability of detection for θ, or the power of the test, is

$$\beta(\theta) = 1 - \mathcal{L}(\theta) \tag{4.17}$$

We may specify one value of $\beta(\theta)$ at a parameter point $\theta = \theta_d$ to yield the bounds $\mathcal{A}$ and $\mathcal{B}$. To determine $\mathcal{L}(\theta)$, consider the statistic

$$\left[\frac{L(\theta_1; \mathbf{Z_n})}{L(0; \mathbf{Z_n})}\right]^{h(\theta)} \tag{4.18}$$

where, for every value θ, the value of $h(\theta)$ is determined such that the expected value of Equation (4.18) is 1, as in

$$\int_{-\infty}^{\infty} \left[\frac{L(\theta_1; \mathbf{Z_n})}{L(0; \mathbf{Z_n})}\right]^{h(\theta)} L(\theta; \mathbf{Z_n})\, \mathrm{d}\mathbf{Z_n} = 1 \tag{4.19}$$

which was derived by Wald [8]. It follows that the integrand of Equation (4.19) is a distribution of $\mathbf{Z_n}$ which we denote by

$$f^*(\mathbf{Z_n}) = \left[\frac{L(\theta_1; \mathbf{Z_n})}{L(0; \mathbf{Z_n})}\right]^{h(\theta)} L(\theta; \mathbf{Z_n}) \tag{4.20}$$

Let H denote the hypothesis that

$$L(\theta; \mathbf{Z_n}) = f(Z_1, \theta) f(Z_2, \theta) \cdots f(Z_n, \theta) = f(\mathbf{Z_n}, \theta) \tag{4.21}$$

is the true distribution of $\mathbf{Z_n}$, and H^* the hypothesis that $f^*(\mathbf{Z_n})$ of Equation (4.20) is the distribution of $\mathbf{Z_n}$. Consider a SPRT which continues taking observations when

$$\mathcal{B}^{h(\theta)} < \frac{f^*(\mathbf{Z_n})}{f(\mathbf{Z_n}, \theta)} < \mathcal{A}^{h(\theta)} \tag{4.22}$$

accepts H^* when the ratio is greater than or equal to $\mathcal{A}^{h(\theta)}$ and accepts H when the ratio is less than or equal to $\mathcal{B}^{h(\theta)}$. From Equation (4.20),

$$\frac{f^*(\mathbf{Z_n})}{f(\mathbf{Z_n}, \theta)} = \left[\frac{L(\theta_1; \mathbf{Z_n})}{L(0; \mathbf{Z_n})}\right]^{h(\theta)} \tag{4.23}$$

so that Equation (4.22) can be rewritten as

$$\mathcal{B} < \frac{L(\theta_1; \mathbf{Z_n})}{L(0; \mathbf{Z_n})} < \mathcal{A} \tag{4.24}$$

which is the standard SPRT derived in Section 4.1.

If the test between H^* and H results in the acceptance of H^*, then Equation (4.24) implies the acceptance of H_1. Likewise, the acceptance of H implies H_0. It follows that $\mathcal{L}(0)$, the probability of accepting H_0, given $\theta = 0$, is the same as $\mathcal{L}(\theta)$, the probability of accepting H when $f^*(\mathbf{Z_n})$ is the true distribution. To calculate $\mathcal{L}(\theta)$, let α' and β' be the size and power of the test of H^* versus H, respectively. It follows that

$$\mathcal{A}^{h(\theta)} = \frac{\beta'}{\alpha'} \tag{4.25}$$

and

$$\mathcal{B}^{h(\theta)} = \frac{1-\beta'}{1-\alpha'} \tag{4.26}$$

Solving for α' and noting that

$$\mathcal{L}(\theta) = 1 - \alpha' \tag{4.27}$$

the operating characteristic function (OCF) is given by

$$\mathcal{L}(\theta) = \frac{\mathcal{A}^{h(\theta)} - 1}{\mathcal{A}^{h(\theta)} - \mathcal{B}^{h(\theta)}} \tag{4.28}$$

when we neglect excess over boundaries as before.

4.4 Average Sample Number

Given the OCF derived in Section 4.3, we can use that information to determine the average number of samples required for detection depending on a particular implementation of the SPRT. Let

$$\mathcal{Z}_n = \log l(\theta; \mathbf{Z_n}) = z_1 + z_2 + \cdots + z_n \tag{4.29}$$

as in the log-likelihood formulation of the SPRT in Section 4.1. The test procedure is as usual for a SPRT: reject H_0 when $\mathcal{Z}_n \geq \log \mathcal{A}$, accept H_0 when $\mathcal{Z}_n \leq \log \mathcal{B}$, and continue the test when $\log \mathcal{B} < \mathcal{Z}_n < \log \mathcal{A}$. If we calculate the average value of $\mathcal{Z}_n$ at a test that has terminated at length n, and we neglect excess over boundaries, then the expectation is given approximately by $\log \mathcal{B}$ times the probability of accepting H_0 plus $\log \mathcal{A}$ times the probability of rejecting H_0,

$$E(\mathcal{Z}_n, \theta) = \mathcal{L}(\theta) \log \mathcal{B} + [1 - \mathcal{L}(\theta)] \log \mathcal{A} \tag{4.30}$$

From the equation for $\mathcal{Z}_n$,

$$E(\mathcal{Z}_n) = E(z_1 + z_2 + \cdots + z_n) = E(n)E(z_i) \tag{4.31}$$

for independent z_i, from $i = 1, \ldots, n$, giving $E(z) = E(z_1) = E(z_2) = \cdots = E(z_n)$. Thus, we denote the average sample number (ASN), $\bar{n} = E(n)$ as

$$\bar{n} = E(n) = \frac{\mathcal{L}(\theta) \log \mathcal{B} + [1 - \mathcal{L}(\theta)] \log \mathcal{A}}{E(z)} \tag{4.32}$$

Given the OCF of a particular sequential test, we can use that to find the average number of samples that will be required for a particular decision, either the acceptance or rejection of H_0. This will be useful to us later in the design of sequential tests to achieve the desired performance, and as a way of characterizing their effectiveness versus other sequential tests or FSS tests.

4.5 Cumulative Sum Tests

Suppose that we are only interested in a one-sided hypothesis test; for example, we only wish to detect the acceptance of H_1 regardless of the times when H_0 is accepted. Or, consider the possibility that we want a test for multiple hypotheses. In these cases, one implementation of the SPRT becomes useful, the cumulative sum test, or CUSUM test. The CUSUM was originally discovered during manufacturing processes when the quality of the output slowly drifted with time, and FSS estimators were unable to detect the slow change. The CUSUM relies on the properties of the logarithm in the log form of the SPRT in that for a single point that is more likely noise than signal, the likelihood ratio will be less than one, resulting in a negative log. For the contrary case, where the point is more likely to come from signal than noise, the likelihood ratio is positive. Thus, for periods in which noise is the predominating factor, the slope of a cumulative sum of each likelihood ratio point

$$\mathcal{Z}_n = z_1 + z_2 + \cdots + z_n \tag{4.33}$$

will be negative, and a positive slope will result from a signal present. Thus, let us only concern ourselves with the bound associated with rejection of H_0, namely when

$$\mathcal{Z}_n \geq \log \mathcal{A} \tag{4.34}$$

We denote the cumulative sum by the notation CUSUM(k,i) where

$$CUSUM(i,k) = \sum_{t=i}^{k} z_t = \mathcal{Z}_k - \mathcal{Z}_i \tag{4.35}$$

In the case of the cumulative sum, there is never an acceptance of H_0; the test merely continues until there is some rejection of H_0. To determine the stopping time of this statistic, we cite the work of Lorden [10]. Lorden determined that any difference between the current CUSUM value and any previous value that exceeds Wald's threshold $\log \mathcal{A}$ would result in the same performance as that of the standard SPRT. Thus, the stopping point is the first point such that

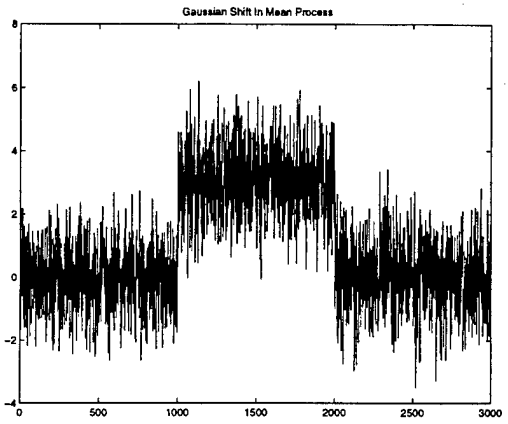
$$n_1 = \inf \left[k \geq 1 : \max_{1 \leq i \leq k} CUSUM(i,k) \geq \log \mathcal{A} \right] \tag{4.36}$$

Using this method, we achieve the same results as the standard SPRT, but in some cases it results in a simplified implementation, when the acceptance of H_0 is not important. As an example, in Figure 4.3(a),(b) we show the CUSUM as it would be formulated for the detection of a shift in the mean of a Gaussian process. We insert a shift in mean from the 1000th to the 2000th sample. During that interval, the CUSUM generates a positive slope. In Figure 4.3(c), we show graphically the interpretation of the stopping rule of Lorden, Equation (4.36).

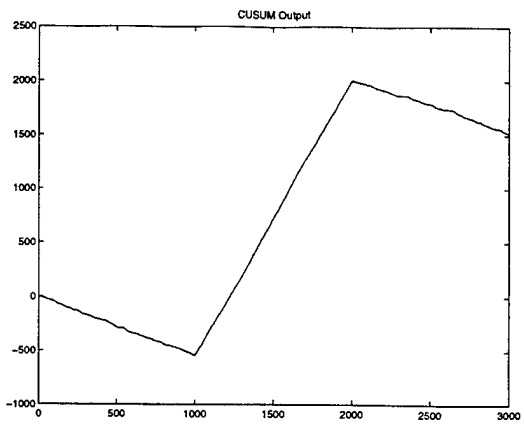
One other advantage of the CUSUM rule is that we can use it as a multihypothesis testing statistic. Consider the problem of estimating the time of arrival (TOA) of a pulse as well as the time of departure (TOD) or end of the pulse. In the standard SPRT, there is no good way of knowing the time of return to the previous statistics. However, for the CUSUM test, we can find the bound easily, by looking for a negative slope that satisfies Wald's bounds. Thus, as in Figure 4.3(c), we find the first point such that

$$n_2 = \inf \left[t \geq N_1 : \max_{N_1 \leq i \leq k} CUSUM(i,k) \leq -\log \mathcal{A} \right] \tag{4.37}$$

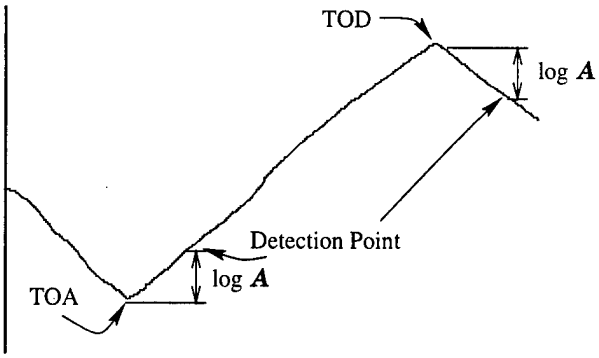
The cumulative sum test is important for ELINT applications as it gives us both the beginning and end of the pulse with similar error bounds. It also alleviates the problem of the SPRT alarming multiple times within a pulse. Since the SPRT resets itself at the acceptance or rejection of H_0, in the duration of a long pulse, we may expect the SPRT to alarm multiple times, which causes needless extra complexity to other processes which rely on accurate TOA times.



(a) Gaussian Shift in Mean Process



(b) CUSUM



(c) Stopping Rules

Figure 4.3 (a),(b) Example of CUSUM output for a Gaussian shift in mean process from the 1000th to the 2000th sample. (c) Graphical interpretation of the stopping rules.

4.6 Optimal SPRT Tests for Detection without Interference

In this and all future sections, we will address the SPRT in general, but the results are applicable as well to the cumulative sum implementation of the SPRT, and the performance characteristics in terms of detection probability and average sample size are identical. We now address the problem of finding a SPRT test suitable to a particular set of statistics. This involves mainly the derivation of the appropriate log-likelihood ratio statistic, and setting the power of the test at a certain parameter point θ_d so that the test achieves desirable results on average.

4.6.1 Gaussian SPRT

As an example that will be influential later in this paper, let us derive the log-likelihood ratio statistic in the case of a Gaussian process where we wish to detect a shift in the mean, assuming the variance of the noise to be constant. The likelihood ratio test for a Gaussian process is of the form

$$l(Z_i;\theta_1,\theta_0) = \frac{\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{1}{2}\frac{(Z_i-\theta_1)^2}{\sigma^2}}}{\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{1}{2}\frac{(Z_i-\theta_0)^2}{\sigma^2}}} \tag{4.38}$$

where the log-likelihood is given by

$$z_i = \log l(Z_i;\theta_1,\theta_0) = \frac{\phi}{\sigma^2}\left(Z_i - \theta_0 - \frac{\phi}{2}\right) \tag{4.39}$$

where $\phi = \theta_1 - \theta_0$ and after a series of algebraic manipulations. The test is defined as the sum of these variables, $\mathcal{Z}_n = z_1 + z_2 + \cdots + z_n$, as in

$$\mathcal{Z}_n = \frac{\phi}{\sigma^2}\sum_{i=1}^{n}\left(Z_i - \theta_0 - \frac{\phi}{2}\right) \tag{4.40}$$

This test is particularly attractive to FPGA implementation due to its ease of implementation, requiring only an estimate of the variance and the mean, as well as additions and two multiplications. This results in relatively little complexity for implementation.

4.6.2 Chi-Squared SPRT

From the section on the polyphase filter bank, we know that our actual statistics are chi-squared random variables. In order to implement the optimal sequential test, we must first derive the log-likelihood ratio as in the case of the Gaussian test. Thus, for a parameter θ_d which in actual cases is the noncentrality parameter λ_d, we find that the likelihood ratio is of the form

$$l(Z_i;\lambda_d) = \frac{\sigma^2\chi_2'^2(\lambda_d)}{\sigma^2\chi_2^2} \tag{4.41}$$

For this derivation we adopt the noncentral chi-square distribution formula developed by Fisher [11]

$$p(x,\lambda) = \frac{1}{2}e^{-\frac{x+\lambda}{2}}I_0(\sqrt{\lambda x}) \tag{4.42}$$

where $I_0(x)$ is the modified Bessel function of the zeroth order. Taking the logarithm and canceling terms, we get

$$z_i = -\frac{\lambda_d}{2} + \log\, I_0(\sqrt{\lambda_d Z_i}) \tag{4.43}$$

This SPRT statistic is optimal in the sense of fitting the expected statistics of the output of the polyphase filter bank. However, for FPGA implementation, the results are nearly impossible to implement because of both the logarithm and the modified Bessel function. Both of these would have to be stored in lookup tables that are expensive in resources to implement. However, let us look at reasonable approximations to the optimal solution. Suppose we expand the logarithmic Bessel function in a power series contingent upon a weak signal in which $\lambda << 1$, which gives us

$$z_i = -\frac{\lambda_d}{2} + \frac{\lambda_d}{4} Z_i - \frac{\lambda_d^2}{64} Z_i^2 + O(\lambda_d^3 Z_i^3) \tag{4.44}$$

It may be noticed that the third term in Equation (4.44) contributes to the bias of z_i, and without that contribution the average sample number approaches infinity for hypothesis H_0. Keeping the terms up to second order results in the need to square each output value. This is also unacceptable in FPGA implementation because of the excessive number of bits that are needed for the computation. A conservative estimate on the output of the filter bank would be 16 bits for an accurate representation of the dynamic range of the input to the ADC. Implementing this form of the SPRT would require us to nominally keep 33 bits for the SPRT, which is far too many to provide efficient implementation. The assumption that $\lambda << 1$ is actually a very reasonable approximation for FPGAs, as a common method is to use fractional two's-complement representation internally. In this method, the numbers are stored in a fractional context with the maximum absolute value of 1. Thus, any noncentrality parameter that does not overflow the ADC would necessarily be less than 1, and much less than 1 in the case of all but the highest SNR pulses where the approximation need not be so accurate as those pulses are easy to detect.

Let us continue approximations that may yield a useful result. From the properties of the chi-squared distribution it can be shown that

$$E(Z_i) = 2 + \lambda \tag{4.45}$$

and

$$E(Z_i^2) = 8 + 8\lambda + \lambda^2 \tag{4.46}$$

Thus, for a received signal with noncentrality parameter λ, the expected value of the test statistic is

$$E(z_i|\lambda) \approx \frac{\lambda_d \lambda}{4} - \frac{\lambda_d^2}{8} \tag{4.47}$$

when terms higher than second order are ignored due to the negligible effects when $\lambda_d << 1$ and $\lambda << 1$. The second term in the approximate expectation is the average bias contribution of the squared Z_i term in the small signal expansion. Suppose now that we keep only the first two terms of the small signal power series and the average contribution of the bias of the third term. It may then be shown that the test will terminate with probability 1; however, average sample numbers with the approximate test may be longer than the optimal implementation using the true bias term or even an exact modified Bessel function. We now have that

$$z_i \approx -\frac{\lambda_d}{2} - \frac{\lambda_d^2}{8} + \frac{\lambda_d Z_i}{4} \tag{4.48}$$

For FPGAs, this approximation is attractive because we have made no assumptions that will in general be violated for our standard operating conditions, and we have reduced the complexity of the implementation to a manageable level. But the question is, what does this approximation tell us? Let us denote the mean of a chi-squared random variable with two degrees of freedom, as at the output of the polyphase filter bank in the case of noise, as $\mu_\chi = 2$ and the variance of the same variable as $\sigma_\chi^2 = 4$. Rewriting Equation (4.48) in an equivalent form we have

$$z_i = \frac{\lambda_d}{\sigma_\chi^2}\left(Z_i - \mu_\chi - \frac{\lambda_d}{2}\right) \tag{4.49}$$

which is exactly equivalent to the Gaussian SPRT statistic of Equation (4.39). Thus, through the small signal approximations we have concluded that the limiting distribution of this set of statistics is Gaussian. This is an important conclusion, as we have derived an optimal small signal chi-squared detector that is equivalent in performance to a Gaussian detector as would be implemented in a coherent detector where the phase of the signal is known. As such, we expect that the test will achieve results that are comparable to matched filter outputs in the case of small signals and fractional two's-complement binary arithmetic.

To characterize the performance of this test, we now derive the OCF of the implementation of the SPRT in Equation (4.49) with the methods described in Section 4.3. As such, we wish to find $h(\lambda)$ such that

$$\int_{-\infty}^{\infty} \left[\frac{L(\lambda_d; \mathbf{Z_n})}{L(0; \mathbf{Z_n})}\right]^{h(\lambda)} L(\lambda; \mathbf{Z_n})\, d\mathbf{Z_n} = 1 \tag{4.50}$$

Using the small signal approximations above, it can be shown that the solution to this equation is approximately

$$h(\lambda) \approx 1 - 2\frac{\lambda}{\lambda_d} \tag{4.51}$$

which, it is interesting to note, is the same $h(\theta)$ that can be derived from the Gaussian SPRT [12]. Thus the OCF for this test is given by

$$\mathcal{L}(\lambda) = \frac{\mathcal{A}^{1-2\lambda/\lambda_d} - 1}{\mathcal{A}^{1-2\lambda/\lambda_d} - \mathcal{B}^{1-2\lambda/\lambda_d}} \tag{4.52}$$

With the OCF, we can also derive the average sample number using the expected value of z_i from Equation (4.47). As in Equation (4.32),

$$\bar{n} = \frac{\mathcal{L}(\lambda)\log\mathcal{B} + [1 - \mathcal{L}(\lambda)]\log\mathcal{A}}{\frac{\lambda_d\lambda}{4} - \frac{\lambda_d^2}{8}} \tag{4.53}$$

4.7 Performance and Design of SPRT Tests

For general SPRT tests, we have so far shown that the performance is based both on the critical bounds $\mathcal{A}$ and $\mathcal{B}$ as well as the design parameter θ_d. In the case of the chi-squared SPRT, the design SNR, represented by the design noncentrality parameter λ_d, is an important consideration, as well as assigning the power of the test at that SNR. Usually, the probability of false alarm is fixed in any application by practical considerations of data volume. In this project, the design value was a FAR of 10^{-3} and we will use that value for subsequent calculations, although the design methods apply to any FAR. It can be seen from the section on determining the critical constants that one of the advantages of the SPRT is the ability to design an arbitrarily powerful test at any design point. We can pick any β at any design point λ_d without restriction except that $0 < \beta < 1$.

However, there are several considerations that the designer should bear in mind when constructing these tests. High power intuitively means that we are accepting a high average sample number as a trade-off. In fact, it will be seen that the ASN of a SPRT is highly nonlinear, and

in some respects can only be a component of designing tests by trial and error. Let us outline a general method of design and provide some examples that will give insight into the important considerations when designing these tests.

In practical applications, it is often important that we achieve detection in as little time as possible, and it has been shown by Wald that the SPRT minimizes the time to detection of any test, when the optimal likelihood ratio is used [8]. Often, we must sacrifice test power to achieve a reasonable detection time. In the case of FSS methods, it was shown that overestimating the observation interval resulted in reduced SNR. In the SPRT, the effect is similar, except that the effect is not on SNR, but on time to detection. If the time to detection is on the average too long, then we run the risk of coming to the end of a pulse before detection. This problem motivates the first step in SPRT test design.

Let us define a general SPRT test S that is implemented in the standard method. For each β on the range from $0 < \beta < 1$ we can calculate the average sample number associated with a received signal at the design parameter λ_d. For increasing power, we expect increasing average sample numbers. Let us define an average sample number n_d which is the desired time to detection for a pulse at SNR commensurate with the design point. From the curve of power versus average sample number, we can read off the maximum power that can be achieved at the design ASN of n_d. To verify that system performance is adequate, we compute the OCF and ASN for the SPRT test S given the power found from the design ASN, and verify that the system responds appropriately at all other SNRs.

Consider the example of the chi-squared SPRT with a design parameter of -5-dB and a design ASN of 8. In other words, we would like to design a test that alarms on the output of the polyphase filter bank and achieves on average a detection time of 8 samples for a -5-dB wideband SNR pulse. This is a reasonable design goal, as it amounts to a detection of a -5-dB pulse that is 2.5 μs or greater. We compute the ASN as a function of test power and note, from Figure 4.4(a), that we can generate a test power of .61 for a design ASN of 8. If we then compute the OCF and ASN of this test based upon a power of .61 and a FAR of 10^{-3} at a design SNR of -5-dB, we can verify that the performance of the test at all other SNRs is reasonable. From Figure 4.4(c), we see that the average sample number peaks at slightly less than -5-dB at approximately 9 samples, which is not unreasonable at lower SNR. Thus we may conclude that this design meets the specifications we imposed at the outset.

Now, let us consider another design process where we set the design SNR slightly lower, at -10-dB and a design ASN of 20. If we again compute the ASN as a function of test power, we see in Figure 4.5(a) that we can achieve a maximum power of approximately .26 for these design constraints. If we now plot the probability of detection derived from the OCF and the ASN associated with the test, we can see in Figure 4.5(c) that the design provides an ASN of less than 20 at all other SNRs, which seems reasonable. However, let us compare the performance of this design at -5-dB with the design previously where the design parameter was -5-dB. Looking at the probability of detection for the second design in Figure 4.5(b), we see that at -5-dB we achieve only a power of .78 at the expense of an ASN that is nearly twice that of the first design. If we set out with an ASN of 15 in the first design, we could achieve greater than a 95% probability of detection. Thus, for a lower SNR design parameter we have sacrificed heavily the optimal performance that could have been achieved at higher SNRs.

This example illustrates an important point. Although we may achieve the design goals of a particular design ASN and design SNR, we may be unduly sacrificing performance at other SNRs if we choose incorrectly the design parameters. It is important to design SPRT tests of this sort through trial and error to achieve acceptable performance at all SNRs. In general there is no optimal way to design these tests so that all SNRs are arbitrarily powerful with acceptable ASNs. Therefore, use of these design methods with a range of design parameters is useful to pick a design parameter that is most favorable in performance over the entire detection range.

4.8 Optimal SPRT Tests for Detection with Interference

As in Chapter 3, we consider the detection of pulses in the presence of CW interference. We saw in the last chapter that in the presence of interference the noise pdf assumes a noncentral chi-square distribution, giving the likelihood ratio as

$$l(Z_i;\lambda_d) = \frac{\sigma^2\chi_2'^2(\lambda_d)}{\sigma^2\chi_2'^2(\lambda_0)} \tag{4.54}$$

where λ_d is the design parameter, and λ_0 is the noncentrality parameter of the interference. If we use Fisher's formula for the $\chi_2'^2(\lambda)$ pdf as we did in the first derivation, it can be shown that the optimal log-likelihood SPRT test is given by

$$z_i = -\frac{\lambda_d - \lambda_0}{2}\log\left(\frac{I_0(\sqrt{\lambda_d Z_i})}{I_0(\sqrt{\lambda_0 Z_i})}\right) \tag{4.55}$$

where $I_0(x)$ is again the modified Bessel function of order 0. Let us apply the same small-signal approximations to $I_0(x)$ as before, keeping the first two terms of the power series expansion and the average bias contribution of the third term. We can then show that the small-signal approximation yields

$$z_i = \frac{\lambda_d - \lambda_0}{4}\left(Z_i - (\lambda_0 + 2) - \frac{\lambda_d - \lambda_0}{2}\right) \tag{4.56}$$

which is exactly the Gaussian SPRT with the mean of the noise correctly given as $\mu_{\chi'} = 2 + \lambda_0$ and the variance as $\sigma^2_{\chi'} = 4$. Thus, we may conclude that the Gaussian approximation is equally valid in the case of CW interference and we need not change critical bounds in this case to achieve similar performance.

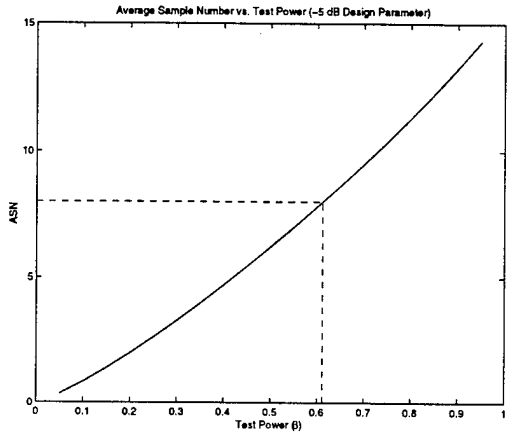
4.9 Comparison of Performance of SPRT versus FSS Methods

Suppose that we want to characterize the performance of the SPRT methods presented in this chapter with the FSS methods presented in the last. Let us first consider the average sample size to detection, and see what the trade-offs between each implementation are. We begin by assuming that we are using a 32-sample window for $\mathbf{Z}_{\mathbf{N}_1}$ and a 64-sample observation interval for $\mathbf{Z}_{\mathbf{N}_0}$. We compute the probability of detection using the F-test for detection of signals with out CW interference. If we then compute the ASN for an SPRT that is designed to achieve the same detection probability curve, we can compare the ASN of the SPRT to detection versus the 32 points required for the FSS test. In Figure 4.6 we see that, while the FSS methods require 32 samples to achieve the same detection probability, the SPRT requires only a maximum of 7 and in some cases as little as 1 sample.

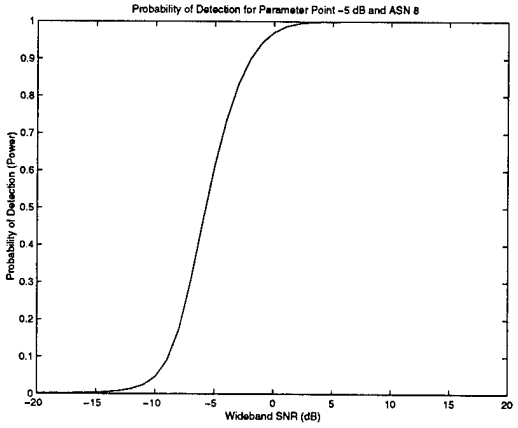
Now, for a similar sample number, let us compare detection probability. Consider the two examples presented in Section 4.7, in which we designed an SPRT based upon an ASN design value and a design SNR. If we design an SPRT in this fashion, and compare its detection probability with an FSS test with sample size equal to the maximum ASN of the SPRT, what is the difference in detection probability? In Figure 4.7 we see that in most cases the SPRT outperforms the FSS tests significantly, except for a small region in the SPRT designed for -10-dB. Thus, we conclude that in all cases, we can design an SPRT that will be superior to the optimal FSS test in both ASN and probability of detection.

4.10 FPGA Implementation

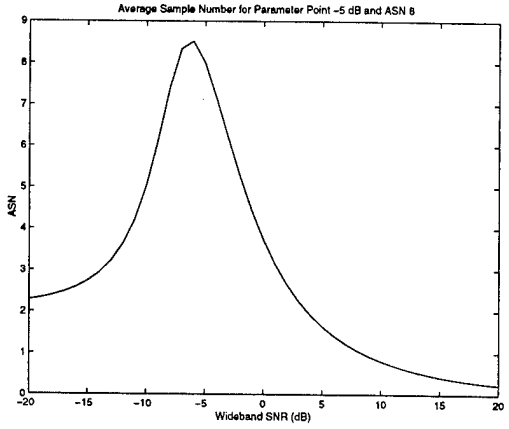
Under the small-signal assumptions, the chi-square SPRT is easily implemented on FPGAs. If we replace $\mu_{\chi'}$ in the CW interference derivation of the optimal SPRT, Equation 4.56, with its ML estimate, the sample mean, we can implement one detector for both interference and noninterference. Likewise, we can replace the variance with an estimate of the variance from the data to account for other inconsistencies in the true distribution of the observed data. In this way, we construct a robust detector that is capable of achieving desired performance in a range of adverse conditions, including interference, nonstationarity, and fixed-precision effects. In addition, this detector fits the statistics well with no unreasonable approximations and results in a low complexity implementation. Thus, we conclude that for ELINT applications using subband channelization based on the DFT, the SPRT is a superior choice for detection over FSS methods.



(a) ASN vs. Test Power

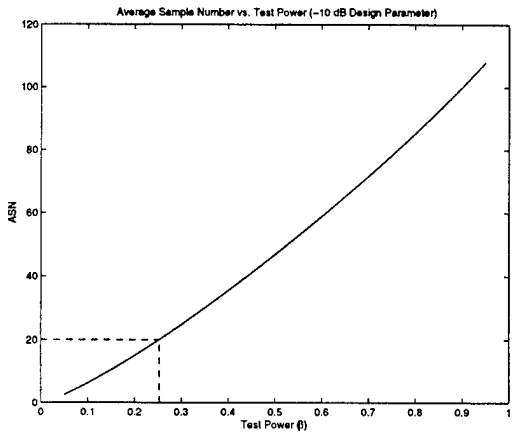


(b) Probability of Detection

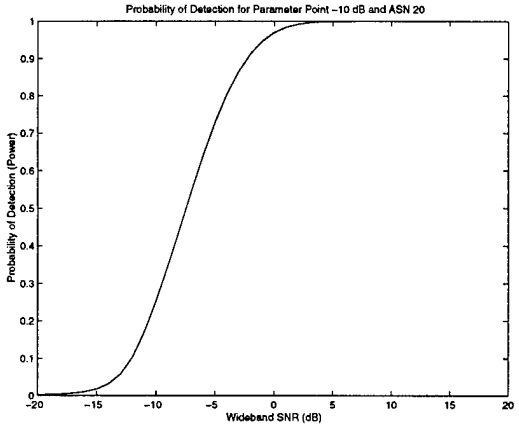


(c) ASN

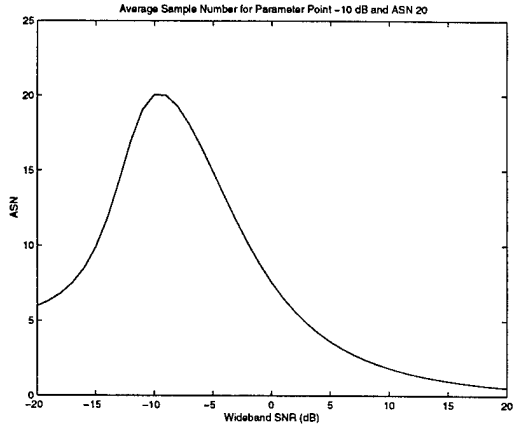
Figure 4.4 (a) ASN of the SPRT test as a function of test power at a design parameter of -5-dB SNR. The dotted lines show the maximum test power that can be achieved at a design ASN of 8. (b) Probability of detection and (c) ASN for a design parameter of -5 dB and a design ASN of 8.



(a) ASN vs. Test Power

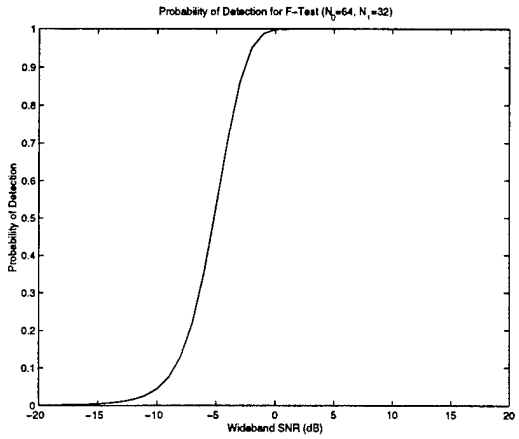


(b) Probability of Detection

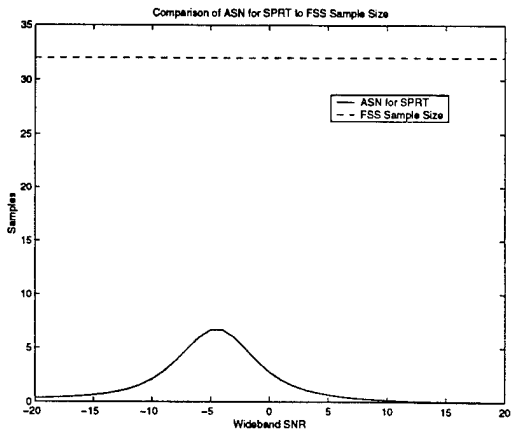


(c) ASN

Figure 4.5 (a) ASN of the SPRT test as a function of test power at a design parameter of -10-dB SNR. The dotted lines show the maximum test power that can be achieved at a design ASN of 20. (b) Probability of detection and (c) ASN for a design parameter of -10-dB and a design ASN of 20.

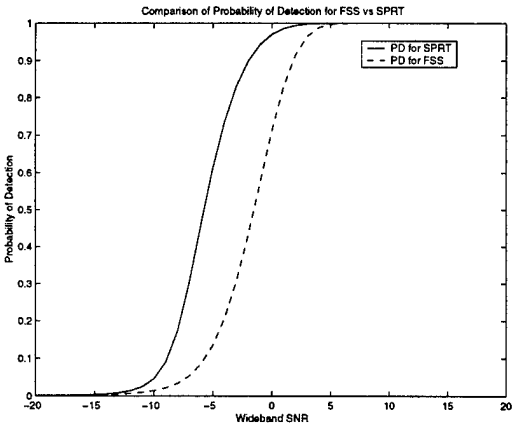


(a) Probability of Detection for both FSS and SPRT Tests

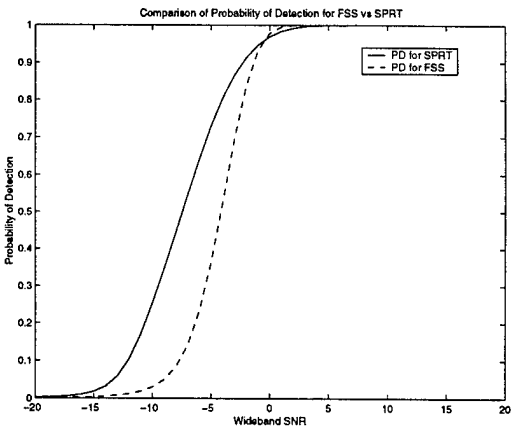


(b) ASN and FSS Sample Size

Figure 4.6 (a) Probability of detection for an FSS test of observations $N_0 = 64$ and $N_1 = 32$. The SPRT is designed to achieve the same detection probability. (b) ASN of the SPRT and sample size of the FSS test to achieve the same detection probability.



(a) -5-dB Design SNR and 8 Sample Design ASN



(b) -10-dB Design SNR and 20 Sample Design ASN

Figure 4.7 (a) Probability of detection for SPRT and FSS methods based on a design SNR of -5-dB and design ASN of 8 for the SPRT, and a FSS sample size of 9. (b)(a) Probability of detection for SPRT and FSS methods based on a design SNR of -10-dB and design ASN of 20 for the SPRT, and a FSS sample size of 21.

CHAPTER 5

CENTER FREQUENCY AND BANDWIDTH ESTIMATION

As a further requirement of this project, we are charged with determining the bandwidth and center frequency of a detected pulse for output to further processing elements that will digitally modulate the signal to baseband and adaptively decimate the signal to the matche' bandwidth estimate of the modulation. If possible, we would like to exploit our estimators used for detection to implement bandwidth calculations.

There are important design goals and limitations to the bandwidth measurement scheme. We desire to provide the system with dehop/dechirp functionality. In other words, we would like to follow the center frequency of the pulse in the case that it is linearly modulated as in the case of LFMOP, or in the case that it takes randomly assigned frequencies as in the case of a hopped NOMOP signal. The limitation imposed upon this functionality is that we can only update the digital baseband modulator on a 1 μs basis.

5.1 FSS Detection Bandwidth

Given a 1 μs basis and a decimation rate of 32 in this project after the DFT, the best we can do at an integer sample size is 4 subband samples, representing a 1.28 μs update rate in the undecimated data stream. Thus, our goal is to update both the bandwidth and center frequency every four subband points after detection, and make that information available for further processing. For the FSS detection methods of Chapter 3, the bandwidth estimation is a further application of FSS methods. Let us define our TOA detection point as $\mathcal{T}$. We would like to retain only statistically significant energy in each channel at a specified confidence level α_{BW}. For this, it seems reasonable to implement the F-test based on four observations from the pulse against N_0 observations of the noise, just as in the FSS scheme presented in Chapter 3.

Thus we calculate the statistic

$$\mathcal{F} = \frac{\frac{1}{4}\sum_{i=K(x)}^{K(x)+3} Z_i}{\frac{1}{N_0}\sum_{N_0} Z_i} \tag{5.1}$$

where $K(x) = [\mathcal{T}, \mathcal{T}+4, \mathcal{T}+8, \ldots]$ encompassing every fourth point from the TOA to the end of the pulse, and $\sum_{N_0} Z_i$ indicates the sum of all points in the window N_0 for the calculation of the noise background. Then, if the statistic

$$\mathcal{F} \geq F_{8,2N_0;\alpha_{BW}} \tag{5.2}$$

we may conclude that there is statistically significant energy in that channel over 1.28 μs to be included in the significant bandwidth. Similarly, in the case that detection is made in interference, the critical bound is given by the noncentral F test.

5.2 SPRT Detection Bandwidth

In the case that the detector is based upon SPRT methods, as is preferable, we would like to exploit that statistic to perform bandwidth estimation without having to recompute FSS statistics. Under the small-signal model, we have an implementation which takes the current point and subtracts an estimate of the mean along with an additional factor based on the design SNR. If we implement the SPRT in the cumulative sum fashion, we notice that this particular implementation is nothing more than an approximate average short-time Fourier transform (STFT) as a result of the polyphase filter bank. The averaged STFT is referred to as a periodogram and is a common method for reducing the variance of the DFT of a stationary process. Thus, to determine significant bandwidth, we need only see which averaged CUSUM point exceeds the background energy level. Consider the statistic

$$\mathcal{U} = CUSUM(K(x)+3, 1) - CUSUM(K(x), 1) \tag{5.3}$$

where $K(x)$ is the same as for FSS estimation. Over a range of four samples, if the CUSUM detects a more probable signal than noise, the slope will be positive on average, and if it is more likely noise than signal, it will be negative. Thus, a reasonable bandwidth estimation scheme would be to construct these statistics with a critical bound of 0. If

$$\mathcal{U} > 0 \tag{5.4}$$

then we conclude that there is significant energy in that channel. In this case, since we are using a mean estimate for the noncentrality parameter of the interference in the case of detection in interference, it may be necessary to subtract an additional confidence factor from the statistic $\mathcal{U}$ to account for the variance in the estimate of the background noncentrality parameter.

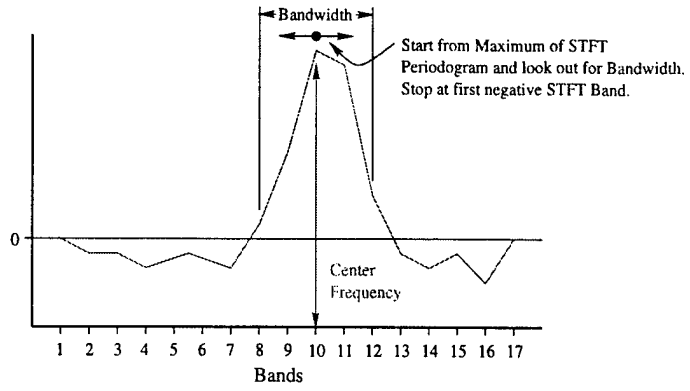
5.3 Advanced Bandwidth Estimation Methods

It is possible to apply advanced methods to the bandwidth estimation problem to determine modulation and other interesting factors about its behavior, but these methods are too complex to implement in today's FPGAs. We provide a note here about some interesting methods that may be possible in the near future when this amount of complexity can be implemented. Under the small-signal model, we can roughly approximate the statistics as being Gaussian. This approximation allows a whole body of theory dealing with the analysis of variance (ANOVA) to be exploited for bandwidth detection. Under these models, we would be able to say, for example, whether there is significant energy in any number of channels simultaneously, or whether interaction between these channels at any stage of detection is significant. This is a powerful method for determining modulation. We can use methods by Scheffe [5] to construct confidence intervals for energy contained between and within channels, as well as models which would explain the interaction over time between channel background noise and the signal energy.

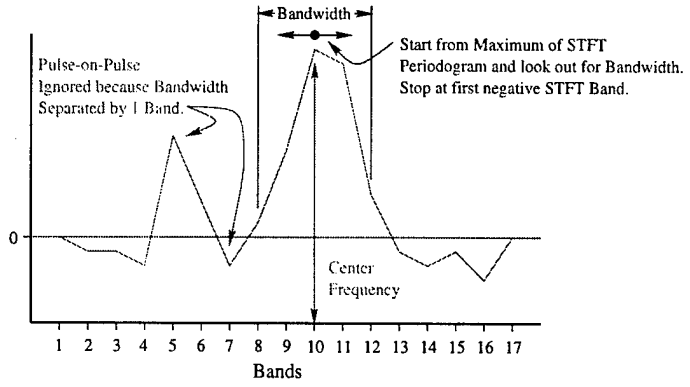
5.4 Excluding Pulse on Pulse Detection

In general, there are two ways of determining the bandwidth of a signal: one would be to take the total bandwidth encompassing any channel with significant energy, and the other would be to take only the bandwidth that is coherently distributed, i.e., occupies adjacent channels. We define a pulse-on-pulse situation as that in which there are two pulses occurring simultaneously with at least one band of separation between their effective bandwidths. In some cases, we would prefer to detect and estimate the bandwidth of only the highest-SNR pulse, eliminating the other through the adaptive filtering process.

To eliminate a pulse-on-pulse bandwidth expansion from occurring, we would like to detect only coherent bandwidth, where the energy is distributed among adjacent bands. Therefore, a logical approach would be to start with the channel of highest energy and look out in both



(a) Normal Pulse Bandwidth



(b) Pulse-on-Pulse Bandwidth

Figure 5.1 Bandwidth calculations excluding pulse-on-pulse occurrences.

directions, towards Nyquist and dc, until the statistic $\mathcal{U}$ has a value less than zero. Thus, in the case that pulses coincide and are separated in bandwidth by at least one band, we can alarm on the highest-SNR signal. We can see in Figure 5.1(b) that the simultaneous pulse is ignored and the bandwidth of the highest-SNR pulse is retained.

5.5 Pulse-on-Pulse Bandwidth Detection Problems

While the pulse-on-pulse bandwidth calculations exclude typical simultaneous pulse matched bandwidth problems, they also cause problems with the bandwidth detection of certain kinds of MOP pulses in which the energy is not coherently distributed. As an example, we show a

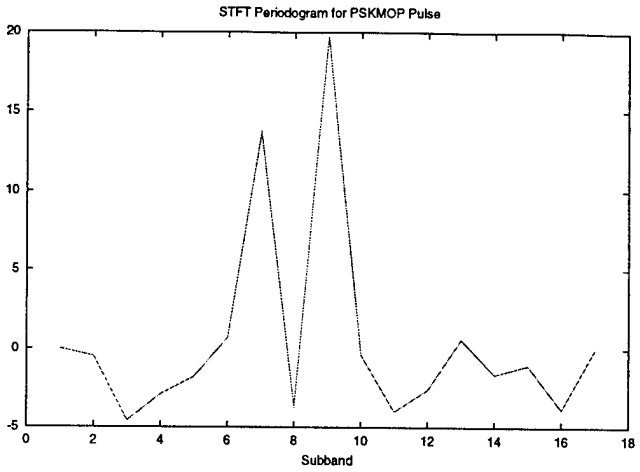


Figure 5.2 STFT from a PSKMOP pulse.

typical periodogram from a PSKMOP pulse. In Figure 5.2 we see that although no pulse-on-pulse behavior is occurring, the modulation bandwidth will not be correctly identified given the noncoherent distribution of energy among the subbands. Thus, we conclude that a better strategy for capturing the true modulation bandwidth is to choose a bandwidth that covers all of the significant energy in any subband, which we call the min/max bandwidth estimation scheme. In Figure 5.3 we use the min/max bandwidth estimation to capture the true bandwidth of the modulation. In the case of a true pulse-on-pulse situation, the result would be the detecting and encompassing of both pulses in the matched bandwidth. Either of the methods discussed for measuring the bandwidth may cause problems depending on the nature of the post-detection processing. Thus, only an application specific choice can be made between the merits of excluding pulse-on-pulse bandwidth at the expense of degraded PSKMOP performance.

5.6 Decimation Rate Stabilization

Because of FPGA implementation and our high input clock rate, the ability to filter the signal to matched bandwidth and then decimate causes important gains in FPGA resource usage as well as the volume of data produced. An adaptive decimator can achieve good performance at matching the pulse bandwidth; however, it is clear that switching the decimation rate in the middle of the pulse will cause unwanted transients and a differing time-frequency scaling

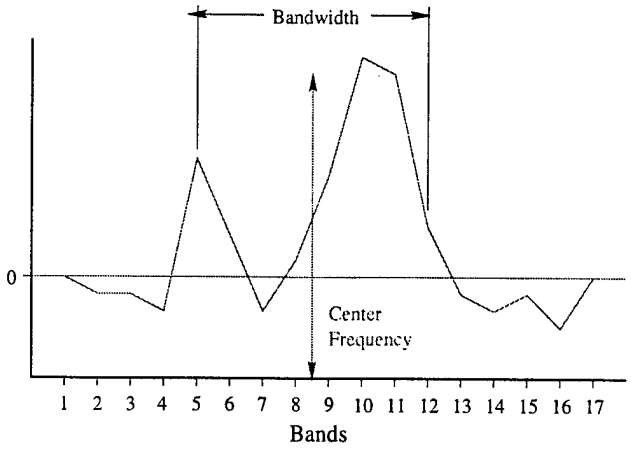


Figure 5.3 Bandwidth calculation using the min/max bandwidth.

throughout the pulse. As such, though we are capable of achieving a 1.28 μs update rate on the tuning of the adaptive decimator, it is unwise to change decimation rates inside a pulse. Thus, we seek some method to calculate a "best" bandwidth that encompasses the maximum amount of energy in the pulse. To do this effectively, the best method would be to take snapshots of the bandwidth of the pulse every 1.28 μs and choose the greatest bandwidth to be sure to capture all the energy over the length of the pulse. Because of FPGA memory limitations, we cannot, in the event of long pulse widths, buffer the entire pulse data. Thus, beginning with the TOA of the pulse $K(1)$, we must calculate a new bandwidth every fourth point using the methods of pulse-on-pulse estimation or min/max estimation. Since there will be some variance in the estimate, we would like to take multiple observations of the bandwidth and choose the largest. We define X_{BW} as the number of bandwidth measurements to be taken before choosing a bandwidth tune word for the set-on receiver. In the event of long pulse detection, as X_{BW} grows, the stability of the bandwidth measurement increases.

5.7 Center Frequency Estimation

After we have calculated our stable bandwidth over the duration of the pulse, what remains for a full tune word to be output to the set-on receiver is the center frequency.

A desirable property of any set-on receiver/detector pair would be the ability to dechirp an LFMOP signal and dehop a NOMOP signal. In the case of this algorithm, we implement this

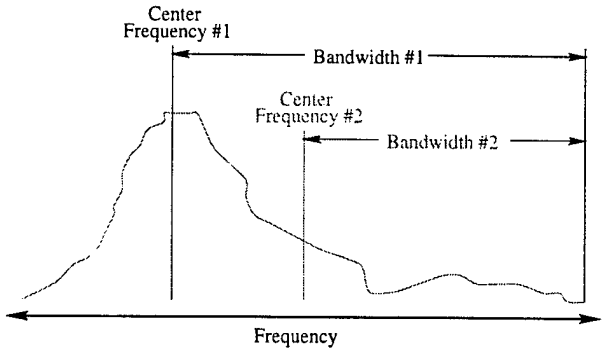


Figure 5.4 Non-baseband-modulated asymmetric bandwidth.

functionality by updating the center frequency on a 1.28 μs basis. Once we have the stable bandwidth from X_{BW} observations of the bandwidth, we return to the TOA point generated by the SPRT and apply that stable bandwidth to finding the center frequency every fourth point. By allowing the center frequency to be updated on a 1.28 μs basis, we provide for the ability to dechirp long-duration LFMOP signals. For short-duration LFMOP, less than the update rate, the full bandwidth is captured, with no dechirping effect. For dehop capability, no center frequency update is needed over a NOMOP pulse as the stable bandwidth should cause the center frequency to remain constant throughout a single pulse. In a pulse train, the separate detection of each pulse will provide a new center frequency in each case, resulting in dehopped output.

5.8 Center Frequency Location

In general, there are two methods for placing the center frequency given a stable bandwidth. One would be to place the center frequency at the highest energy level subband. The second would be to place the center frequency at the middle of the bandwidth that covers the significant energy. To clarify, we offer this example, where there are the two options above. In Figure 5.4, we show a pulse situation with an asymmetric bandwidth, with the two options for center frequency calculation. Here, we assume the ability to adjust the bandwidth to cover the coherent energy spectrum. Choosing center frequency and bandwidth #1, as in Figure 5.5, we see that we introduce noise into the matched bandwidth filtering because of the asymmetric bandwidth. If we choose center frequency #2, as in Figure 5.6, we realize a true matched bandwidth process

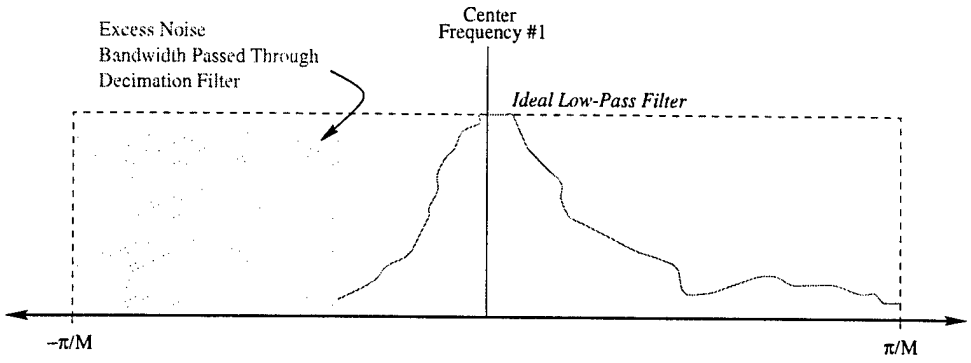


Figure 5.5 Excess noise introduced to matched bandwidth for center frequency #1.

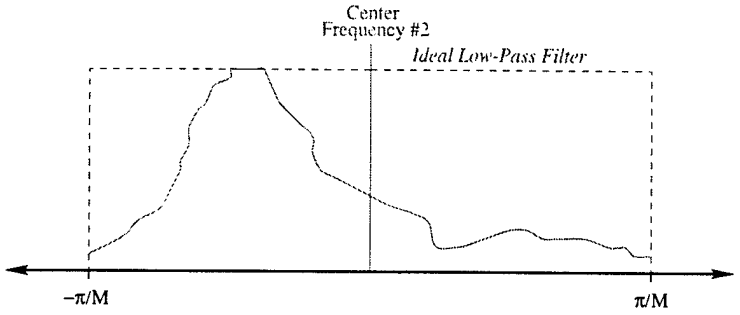


Figure 5.6 True matched bandwidth for center frequency #2.

by optimally filtering the significant bandwidth. The only error introduced in this method would be in the case where the true center frequency is placed on the filter transition band edge, which would result in some attenuation of that center frequency due to nonideal low-pass filtering.

5.9 Center Frequency Calculation

Given the above choices for center frequency, we choose to locate our center frequency at the middle of the stable bandwidth for optimal matched bandwidth filtering. Once we obtain the stable bandwidth estimate, we must return to the TOA location $K(1)$ and begin a center frequency calculation every fourth point. We use a method similar to bandwidth estimation by first calculating the statistic $\mathcal{U}$. After we generate the periodogram, we would like to maximize the energy covered by the stable bandwidth. Given the stable bandwidth and the 27 valid steps of our baseband modulator, we realize some number of allowable center frequencies for

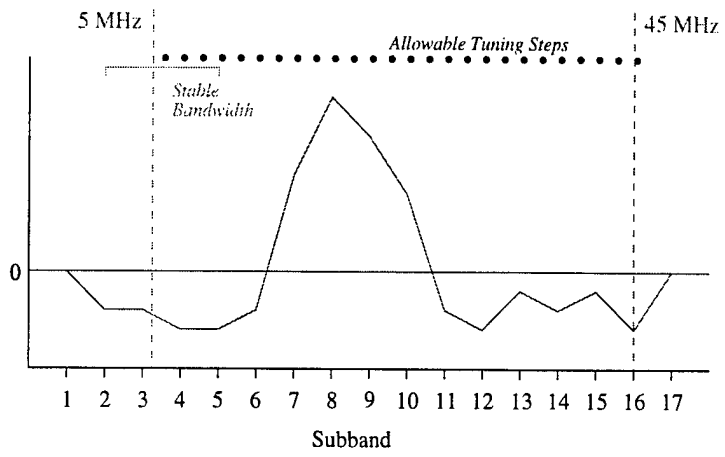


Figure 5.7 Allowable tuning steps for center frequency

the bandwidth to be contained between the dc and Nyquist frequency. By summing the energy contained in the periodogram under the stable bandwidth at each step, we simply choose the center frequency that generates the maximum energy. This results in a center frequency that optimally covers the most significant energy at each 1.28 μs update.

5.10 FPGA Implementation

Using the statistics already derived in either the FSS or SPRT methods provides simple methods of bandwidth and center frequency estimation. Each method involves very little computational complexity. The advanced ANOVA techniques would require significantly more computational complexity, possibly involving a further transformation of the data, and squaring operations. These methods are unsuitable for implementation in this project at this point; however, they may be useful in future applications on more advanced FPGAs or in a floating-point format.

REFERENCES

[1] R. Wiley, *Electronic Intelligence: The Analysis of Radar Signals.* Dedham, MA: Artech House Inc., 1993.

[2] J. L. Arrowood, "Comparison of filter banks for signal detection," *Los Alamos Unclassified Report (LAUR 99-4551)*, 1999.

[3] R. Hogg, *Introduction to Mathematical Statistics.* Upper Saddle River, NJ: Prentice-Hall, Inc., 1995.

[4] N. Johnson and S. Kotz, *Distributions in Statistics: Continuous Univariate Distributions-2.* New York, NY: John Wiley and Sons, 1970.

[5] H. Scheffe, *The Analysis of Variance.* New York, NY: John Wiley and Sons, 1959.

[6] W. Bulgren, "On representations of the doubly non-central F distribution," *JASA*, vol. 66, pp. 184–186, 1971.

[7] M. Tiku, "A note on the distribution of the doubly non-central F distribution," *Austral. J. Statist.*, vol. 14, pp. 37–40, 1972.

[8] A. Wald, *Sequential Analysis.* New York, NY: John Wiley and Sons, 1947.

[9] E. Page, "Continuous inspection schemes," *Biometrika*, vol. 41, pp. 100–115, 1954.

[10] G. Lorden, "Procedures for reacting to a change in distribution," *Annals of Mathematical Statistics*, vol. 42, pp. 1897–1908, 1971.

[11] R. Fisher, "The general sampling distribution of the multiple correlation coefficient," *Proceedings of the Royal Society of London*, vol. 121A, pp. 654–673, 1928.

[12] J. DiFranco and W. Rubin, *Radar Detection.* Dedham, MA: Artech House, Inc., 1980.